
Heterochromatin dynamics and functions

Matthias Ulrich Hahn

Dissertation
der Fakultät für Biologie
der Ludwig-Maximilian-Universität
München

vorgelegt von
Matthias Ulrich Hahn
aus Ulm

München, den 10.01.2013

Erstgutachter: Prof. Dr. Peter Becker

Zweitgutachter: Prof. Dr. Heinrich Leonhardt

Tag der mündlichen Prüfung: 02.05.2013

Statement and statutory declaration

Ich erkläre hiermit, dass die Dissertation keiner anderen Prüfungskommission vorgelegt worden ist und dass ich mich nicht anderweitig einer Doktorprüfung ohne Erfolg unterzogen habe. Ich versichere hiermit an Eides Statt, dass die vorgelegte Dissertation von mir selbständig, ohne unerlaubte Hilfe angefertigt wurde.

Ort, Datum

Unterschrift

Index

Statement and statutory declaration	4
Index	5
Abbreviations	6
List of publications	7
Declaration of contributions	8
Zusammenfassung.....	10
Summary	12
1. Introduction	14
1.1 Epigenetic gene regulation	14
1.2 Pericentric heterochromatin	16
1.3 Chromatin changes in development.....	18
1.4 Polycomb gene silencing via H3K27 methylation.....	19
1.5 Higher order chromatin structure.....	20
1.6 Centromeres and heterochromatin.....	21
1.7 Aims of this work	24
2. Results	25
2.1 Heterochromatin dysregulation in human diseases.	25
2.2 Epigenetic regulation of development by histone lysine methylation.	37
2.3 Suv4-20h histone methyltransferases promote neuroectodermal differentiation by silencing the pluripotency-associated Oct-25 gene.....	52
2.4 H3K56me3 is a novel, conserved heterochromatic mark that largely but not completely overlaps with H3K9me3 in both regulation and localization.	97
2.5 Suv4-20h2 mediates chromatin compaction and is important for cohesin recruitment to heterochromatin.	111
2.6 CENP-C facilitates the recruitment of M18BP1 to centromeric chromatin.	167
3. Discussion	189
3.1 Suv4-20h enzymes and H4K20me3 in embryonic development.....	189
3.2 H3K56me3: a novel pericentric heterochromatin mark	190
3.3 Suv4-20h enzymes in heterochromatin organization	192
3.4 Heterochromatin in centromere formation	197
3.5 Outlook	199
Curriculum Vitae.....	202
Acknowledgements.....	203
References	204

Abbreviations

3D-SIM	3-dimensional structured illumination microscopy
AurB	Aurora B kinase
CCAN	constitutive centromere associated network
ChIP	chromatin-immunoprecipitation
DAPI	4',6-diamidino-2-phenylindole
DKO	double knock-out
EB	embryoid body
ESC	embryonic stem cell
FCS	fluorescence correlation spectroscopy
FACS	fluorescence activated cell sorting
FISH	fluorescence in situ hybridization
FL	full length
FRAP	fluorescence recovery after photobleaching
GST	Glutathion-S-transferase
H1	linker histone H1
HJURP	Holliday junction-recognizing protein
HP1	heterochromatin protein 1
IF	immunofluorescence
IP	immunoprecipitation
K	lysine
KMT	histone lysine methyltransferase
KDM	histone lysine demethylase
KO	knock-out
MBD	methyl binding domain
MEF	mouse embryonic fibroblast
me1	monomethylation
me2	dimethylation
me3	trimethylation
MNase	micrococcal nuclease
ncRNA	noncoding RNA
PcG	polycomb group
qPCR	quantitative polymerase chain reaction
Sgo1	Shugoshin-like 1
TrxG	trithorax
WT	wild type

List of publications

Hahn M*, Dambacher S*, Dulev S, Kuznetsova A, Eck S, Wörz S, Sadic D, Schulte M, Mallm JP, Maiser A, Debs P, Melchner H, Leonhardt H, Rohr K, Schermelleh L, Rippe K, Storchova Z, Schotta G. Suv4-20h2 mediates chromatin compaction and is important for cohesin recruitment to heterochromatin. *Genes & Development*. 2013 Apr 15; 27(8):859-72

Müller-Ott K, Erdel F, Matveeva A, **Hahn M**, Mallm JP, Marth C, Zhang Q, Kaltofen S, Schotta G, Höfer T, Rippe K. Specificity, propagation and epigenetic memory of pericentric heterochromatin in mouse fibroblasts. *manuscript in revision*

Nicetto D, **Hahn M**, Jung J, Schneider T, Straub T, David R, Schotta G and Rupp R. Suv4-20h histone methyltransferases are required for neuroectodermal differentiation and for silencing of pluripotency-associated POU-V genes. *PLOS Genetics*. 2013 Jan;9(1):e1003188.

Jack A, Bussemer S, **Hahn M**, Pünzeler S, Wells M, Solovei I, Csankovszki G, Schotta G, Hake S. H3K56me3 is a novel, conserved heterochromatic mark that largely but not completely overlaps with H3K9me3 in both regulation and localization. *PLOS One*. 2013 Feb;8(2):e51765.

Dambacher S*, Deng W*, **Hahn M***, Sadic D, Fröhlich J, Nuber A, Hoischen C, Diekmann S, Leonhardt H, Schotta G. CENP-C facilitates the recruitment of M18BP1 to centromeric chromatin. *Nucleus*. 2012 Jan 1;3(1):101-10.

Hahn M*, Dambacher S*, Schotta G. Heterochromatin dysregulation in human diseases. *J Appl Physiol*. 2010 Jul;109(1):232-42.

Dambacher S*, **Hahn M***, Schotta G. Epigenetic regulation of development by histone lysine methylation. *Heredity (Edinb)*. 2010 Jul;105(1):24-37.

Rivière L, Moreau P, Allmann S, **Hahn M**, Biran M, Plazolles N, Franconi JM, Boshart M, Bringaud F. Acetate produced in the mitochondrion is the essential precursor for lipid biosynthesis in procyclic trypanosomes. *Proc Natl Acad Sci U S A*. 2009 Aug 4;106(31):12694-9

* joint first authors

Declaration of contributions

Declaration of contributions to “Heterochromatin dysregulation in human diseases.”

This review was conceived by Gunnar Schotta, Silvia Dambacher and me. I was accountable for the introductory part on epigenetic gene regulation, the sections on “Triplet repeat induced heterochromatin formation in Friedreich’s Ataxia” and “Epigenetic dysregulation of imprinting control regions in Prader-Willi Syndrome and Angelman Syndrome”. I created Figure 1, Figure 3 and Figure 4. Gunnar Schotta, Silvia Dambacher and me developed the final manuscript.

Declaration of contributions to “Epigenetic regulation of development by histone lysine methylation.”

This review was conceived by Gunnar Schotta, Silvia Dambacher and me. I was accountable for the sections on “[Gene] Activation and repression are facilitated by histone lysine methylation”, “Repressive histone lysine methylation marks” and “Repression of developmental regulators by H3K27 methylation”. I created Figure 1, Figure 3, Figure 4 and Figure 5. Gunnar Schotta, Silvia Dambacher and me developed the final manuscript.

Declaration of contributions to “Suv4-20h histone methyltransferases promote neuroectodermal differentiation by silencing of pluripotency-associated Oct-25 gene.”

This study was done in collaboration with the lab of Prof. Dr. Ralph Rupp. I conducted the mouse cell transfections and immunofluorescence (IF) experiments. Furthermore, I established and performed the differentiation marker qPCR analysis of embryonic stem cells (ESCs) and embryoid bodies (EBs).

Declaration of contributions to “H3K56me3 is a novel, conserved heterochromatic mark that largely but not completely overlaps with H3K9me3 in both regulation and localization.”

This project was done in collaboration with the group of PD Dr. Sandra Hake. I cloned the jmjC-plasmid-library for the demethylase screen into pEGFP-N1-Gateway vectors. In addition, I performed the H3K56/H3K9 methylation IF analyses in wildtype and Suv39h mutant mouse embryonic fibroblasts (MEFs).

Declaration of contributions to “CENP-C facilitates the recruitment of M18BP1 to centromeric chromatin.”

I performed all cell-cycle-dependent M18BP1 localization IF analyses and CENP-C colocalization IF experiments in mouse ESCs. Further, I conducted the IF analyses of CENP-C knockdowns in ESCs. Gunnar Schotta wrote a first draft of the paper and together with Silvia Dambacher, we wrote the final manuscript.

Declaration of contributions to “Suv4-20h2 mediates chromatin compaction and is important for cohesin recruitment to heterochromatin.”

This study was conceived by Gunnar Schotta, Silvia Dambacher and me. I designed and performed all fluorescence recovery after photobleaching (FRAP) experiments. I cloned pmCherry plasmids for the F2H assay. I cloned parts of the Suv4-20h truncation library for IF and recombinant protein expression experiments. I conducted the mouse major satellite fluorescence in situ hybridization (FISH) experiments, the IF analyses of the HP1 knock-out MEFs and Suv4-20h2 truncation localizations, cell cycle synchronization and FACS (fluorescence assorted cell sorting) analyses. I performed the mitotic spreads, analyzed the mitotic defects and prepared the final figures. Gunnar Schotta wrote a first draft of the paper and together with Silvia Dambacher we wrote the final manuscript.

Prof. Dr. Gunnar Schotta

Silvia Dambacher

Matthias Hahn

Zusammenfassung

Im Zellkern liegt Chromatin als hochorganisierter DNA/Protein-Komplex vor, der dynamisch reguliert werden muß. Kompaktierte DNA mit stillgelegten Genen bildet heterochromatische Regionen aus, und weniger dicht gepackte DNA genetisch aktive euchromatische Regionen. Diese genomischen Bereiche werden von einem Netzwerk verschiedenster epigenetischer Mechanismen reguliert. Diese Steuerung ist unverzichtbar, da Heterochromatin wichtige Rollen in der Zelldifferenzierung, Zellkernarchitektur und Sicherung der genomischen Integrität spielt.

Diese Dissertation besteht aus zwei Reviewartikeln und vier Originalartikeln, die verschiedene Aspekte der Funktion und Regulation von zentromerischen und perizentromerischen Chromatins erläutern. Im Artikel *Nicetto et al. 2013* wird eine neue Funktion der Suv4-20h Histonmethyltransferasen in Entwicklungsprozessen von Vertebraten beschrieben. Suv4-20h Enzyme werden für die Repression des *Xenopus laevis* Oct-25 Gens benötigt. Das XOct-25 Gen ist ein Homolog des Oct4 Gens von Säugetieren, welches als Transkriptionsfaktor wichtige Aufgaben für die Regulation der Pluripotenz von Stammzellen bewerkstelligt. Die Charakterisierung von embryonalen Maus Stammzellen zeigte, dass auch in diesem System Suv4-20h Enzyme an der Transkriptionsregulation von Oct4 beteiligt sind. Der Verlust der Suv4-20h Methyltransferasen führte im Frosch, sowie in den Maus Stammzellen zu Entwicklungsdefekten. Zusammengefasst lassen diese Ergebnisse vermuten, dass Suv4-20h vermittelte H4K20me3 an POU-V Pluripotenzgenen ein konservierter Regulationsmechanismus zur Gewährleistung korrekter Zelldifferenzierungsprozesse ist.

Jack et al. beschreibt die Charakterisierung der neuen Histonmodifikation H3K56me3. Diese ist evolutionär konserviert und wird durch Suv39h Enzyme vermittelt. Darüberhinaus konnte gezeigt werden, dass Proteine aus der Gruppe der JmjD2-Histondemethylasen die H3K56me3 Modifikation entfernen können.

Im Artikel *Hahn et al.* können wir zeigen, das Suv4-20h2 wichtige Rollen in der Regulation der Zellkernarchitektur, Chromosomensegregation und Rekrutierung von Cohesin spielt. Suv4-20h2 kann durch synergistische Interaktionen mit mehreren HP1 Proteinen stabil an perizentromerische Heterochromatinregionen binden und diese Chromatinbereiche kompaktieren. Suv4-20h Knockout Zelllinien wiesen weniger kompaktiertes Heterochromatin und Defekte in der mitotischen Chromosomensegregation auf. Die Untereinheiten des Cohesin Komplex, Smc1 und Smc3, können mit Suv4-20h2 interagieren. Die Analyse von Suv4-20h2 Knock-out

Zelllinien zeigte, dass Suv4-20h2 für die Rekrutierung von Cohesin an das perizentromerische Heterochromatin benötigt wird um, somit die Stabilität des Genoms zu wahren.

Die Zentromerbereiche der Chromosomen werden wahrscheinlich durch epigenetische Mechanismen definiert. Die zentromerspezifische Histonvariante Cenp-A muß ins zentromerische Heterochromatin integriert werden um die Ausbildung eines funktionalen Zentromers zu ermöglichen. Das M18bp1 Protein wird benötigt um Cenp-A Histone ins Chromatin einzubauen. Im Artikel *Dambacher et al. 2012* wird die Assoziation von M18bp1 in Abhängigkeit des Zellzyklus beschrieben. Eine Interaktion von M18bp1 und dem zentromerischen Protein Cenp-C konnte *in vivo* und *in vitro* nachgewiesen werden. Knock-down von Cenp-C führte zu verminderten M18bp1 und Cenp-A Proteinmengen an Zentromeren, was vermuten lässt, dass Cenp-C einen wichtigen Faktor in zentromerischen M18bp1 und Cenp-A Ladeprozessen darstellt.

Summary

Chromatin is forming higher order structures in the nucleus of the cell. This three-dimensional organization, needs to be dynamically coordinated to allow gene regulation. For proper cell function, it is important to control the loosely packed, transcriptionally active euchromatin and the densely packed, transcriptionally silent heterochromatin. Various epigenetic mechanisms are part of a regulatory network that defines and modulates these genomic domains. Heterochromatin and its protein framework, have been shown to play fundamental roles in lineage decision processes, in governing nuclear architecture and guaranteeing genomic integrity.

This thesis consists of two review articles and four published papers dealing with the biological relevance of pericentromeric or centromeric chromatin and proteins.

In *Nicetto et al. 2013*, a novel function of the heterochromatic histone methyltransferase Suv4-20h in vertebrate development is described. A multitude of different gene expression patterns, needs to be tightly orchestrated during embryogenesis. In *Xenopus laevis* cell differentiation, Suv4-20h enzymes are shown to be required for repression of the Oct-25 pluripotency gene, a homolog of mammalian Oct4. Consistently, mouse ESCs lacking Suv4-20h display increased Oct4 protein levels before and during *in vitro* differentiation. As a consequence, aberrant and biased differentiation behaviour can be observed, suggesting that Suv4-20h-mediated H4K20me3 at specific POU-V pluripotency genes may be a conserved regulatory mechanism regulating cell fate decisions.

Jack et al. reports on histone H3 lysine (K) 56me3 (tri-methylation) as a novel and evolutionarily conserved heterochromatin modification. Suv39h enzymes are mediating trimethylation of H3K56 in pericentric heterochromatin and the Jmjd2 family of demethylases is capable to remove this modification.

Higher order chromatin organization needs to be controlled in a complex manner throughout the cell cycle. In *Hahn et al.*, the important roles that Suv4-20h2 is playing in regulating nuclear architecture, ensuring faithful chromosome segregation and cohesin recruitment, are depicted. Moreover, evidence is provided, that Suv4-20h2 stably binds to pericentric heterochromatin via multiple synergistic interactions with HP1 proteins and that Suv4-20h2 is important for the compaction of these regions. Suv4-20h knock-out cells show reduced chromatin compaction, defects in chromocenter organization and errors in chromosome segregation during mitosis. Furthermore, Suv4-20h2 is identified as an interactor of the cohesin subunits Smc1

and Smc3. These results indicate that Suv4-20h2 is necessary for efficient cohesin recruitment to heterochromatin and to ensure genomic integrity.

In higher eukaryotes, the centromere regions of chromosomes are likely determined by epigenetic mechanisms. The histone variant Cenp-A must be integrated into these chromosomal domains to recruit other centromeric proteins that build up a functional kinetochore. The Mis18 complex, and in particular its member M18bp1, are known to be essential for both incorporation and maintenance of Cenp-A. The publication *Dambacher et al. 2012* describes a cell cycle-regulated association of M18bp1 with centromeric chromatin. An interaction with Cenp-C is shown in *in vivo* and *in vitro* experiments. Depletion of Cenp-C results in diminished M18bp1 and Cenp-A levels at centromeres, indicating that Cenp-C works as an important factor for centromeric M18bp1 recruitment and thus for maintaining centromeric Cenp-A deposition.

1. Introduction

Nuclear DNA is packaged with proteins in complex called chromatin. The repeating unit of chromatin is the nucleosome, consisting of ~147bp of DNA wrapped around a histone octamer (two each of histones H2A, H2B, H3, and H4) (Luger et al. 1997). Nucleosomes are bridged by the linker DNA with a variable length of ~20 to 80bp, that can serve as a binding surface for the linker histone H1. These interactions contribute to stabilizing higher order chromatin folding and compaction (Lu et al. 2009a). It is important to precisely control the chromatin states and nuclear architecture. Specific posttranslational modifications of histones, chromatin remodelling, incorporation of histone variants and chemical modifications of the underlying DNA are known to modulate chromatin. A coordinated interplay of these mechanisms is crucial for proper cell function and the establishment of defined gene expression programs in cellular differentiation.

1.1 Epigenetic gene regulation

The following subchapter and Figure 1 have been published in (Hahn et al. 2010), a review article, of which I am a joint first author (for author contributions see page 8). Slight changes in the text (marked in *italics*) have been conducted.

The grade of chromatin compaction can influence DNA accessibility to the transcriptional machinery. Euchromatin is loosely packed and transcriptionally active, whereas heterochromatin represents a more densely packed chromatin state that is characterized by low transcriptional activity. Heterochromatin can be subdivided into two categories: facultative heterochromatin for silencing of developmental genes and the inactive X chromosome in female mammals, and constitutive heterochromatin, which is formed at pericentromeric regions and telomeres.

Five major epigenetic mechanisms are capable of establishing and stabilizing open or closed chromatin structures, thereby regulating transcriptional activity.

1. Nucleosome remodelers can stimulate transcription by removing nucleosomes from promoter regions, allowing transcription factors to gain access to the underlying DNA (Korber et al. 2004). Additional functions of nucleosome remodelers include histone variant exchange and nucleosome sliding (Varga-Weisz and Becker 2006).
2. Mammalian cells express three histone H3 variants: H3.1, H3.2, and H3.3. Specific enrichment of H3.3 in transcriptionally active genes and regulatory regions, and H3.1 in repressed or inactive genetic elements suggest a regulatory function of these variants (Mito et al. 2005). Incorporation of histone variants is also important for many other chromatin-related processes. Faithful DNA damage repair, for

example, depends on the presence of a histone H2A variant, H2A.X, which is phosphorylated on damage detection allowing binding of the DNA damage repair machinery (Celeste et al. 2002).

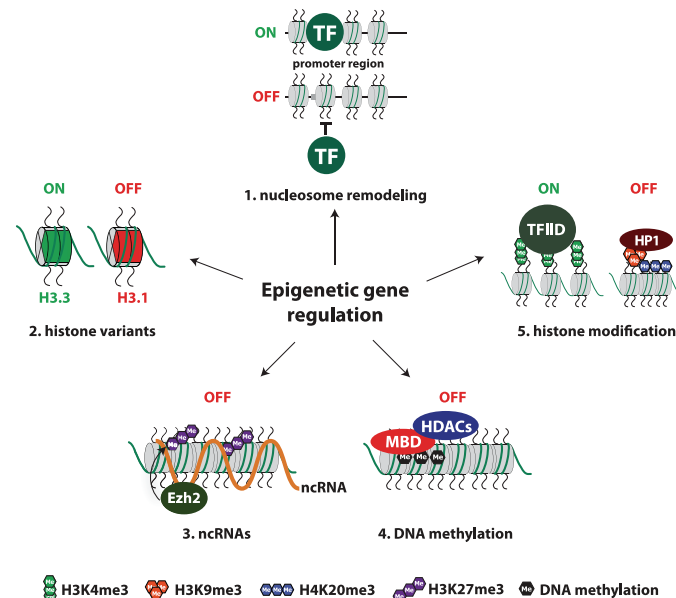


Figure 1. Major epigenetic mechanisms that regulate chromatin structure (Hahn et al. 2010)

1. Nucleosomes can prevent binding of transcription factors (TF). One function of nucleosome remodeling complexes is to evict nucleosomes to make promoters accessible. 2. Distinct histone H3 variants associate with transcriptionally active or repressive chromatin domains. 3. Noncoding RNAs (ncRNAs) can recruit histone-modifying enzymes. For example, interaction of the H3K27-specific histone methyltransferase Ezh2 with a ncRNA is important to induce X inactivation. 4. DNA methylation is a repressive modification that is recognized by specific binding factors (MBD proteins) which recruit other corepressor proteins, e.g., HDACs. 5. Histone post translational modifications (PTM) can recruit binding partners that induce active or repressive chromatin states. H3K4me3 correlates with transcriptional activity and is recognized by components of the TFIID complex. Heterochromatic chromatin structures feature the combinatorial mark H3K9me3 and H4K20me3.

3. Increasing evidence suggests important functions for long noncoding RNAs (ncRNAs) in transcriptional regulation. Direct interactions of chromatin-modifying enzymes with non-coding RNAs are speculated to facilitate targeting to specific genomic loci. A well-studied example is the histone methyltransferase (KMT) Ezh2, which was shown to interact with different ncRNAs to induce X inactivation and repression of developmental genes (Rinn et al. 2007; Zhao et al. 2008). Small ncRNAs, e.g., siRNAs and promoter antisense RNAs, can also trigger formation of repressive chromatin structures.

4. Cytosine bases, preferably in the context of CpG dinucleotides, can be methylated to 5-methylcytosine by DNA methyltransferases. DNA methylation is a repressive modification that is enriched at promoter regions of genes and noncoding DNA

sequences. MBD (methyl binding domain) proteins and MeCP2 can bind methylated DNA stretches and in turn recruit corepressor complexes to facilitate transcriptional silencing (Sasai and Defossez 2009).

5. Histones can be posttranslationally modified. The major modifications include phosphorylation, acetylation, and methylation. Combinations of different histone modifications represent chromatin signals that are recognized by specific binding proteins that then mediate downstream effects. In the context of transcriptional regulation, histone lysine methylation has been particularly well characterized. This modification generates a high complexity of signals as each lysine position can be mono- (me1), di- (me2), or trimethylated (me3) and distinct binding proteins for each methylation state can mediate different functions. Transcriptionally active, euchromatic domains are characterized by histone H3 lysine 4 trimethylation (H3K4me3) at gene promoters and H3K36me3 across gene bodies (Barski et al. 2007). The two types of heterochromatin carry distinct modification patterns. Facultative heterochromatin is marked by high levels of H3K27me3 (Trojer and Reinberg 2007). In contrast, constitutive *also referred to as pericentric* heterochromatin, features the combinatorial mark H3K9me3 and H4K20me3 (Lachner et al. 2004).

1.2 Pericentric heterochromatin

The following subchapter and Figure 2 have been published in (Dambacher et al. 2010), a review article, of which I am joint first author (for author contributions see page 8). Slight changes in the text (marked in *italics*) have been conducted.

The largest family of repetitive regions consists of major satellite repeats that are the main constituents of pericentric heterochromatin. Major satellite repeats have a distinct H3K9me3 and H4K20me3 chromatin signature, which is found in almost all cell types and developmental stages, suggesting that these modifications have a general function in heterochromatin. H3K9me3 is established by Suv39h1 and Suv39h2 enzymes (Rea et al. 2000). Two other HMTases, Suv4-20h1 and Suv4-20h2, establish H4K20me3 (Schotta et al. 2004). Recent data suggest a novel modification, H3K64me3 that has a role in pericentric heterochromatin formation during the early stages of mouse development (Daujat et al. 2009); however, the responsible methyltransferase has not been identified as yet.

The combinatorial pattern of histone lysine methylation at heterochromatin is established in a sequential pathway (Figure 2). In *Schizosaccharomyces pombe*, double-stranded RNA from centromeric repeats is processed by components of the RNA interference machinery.

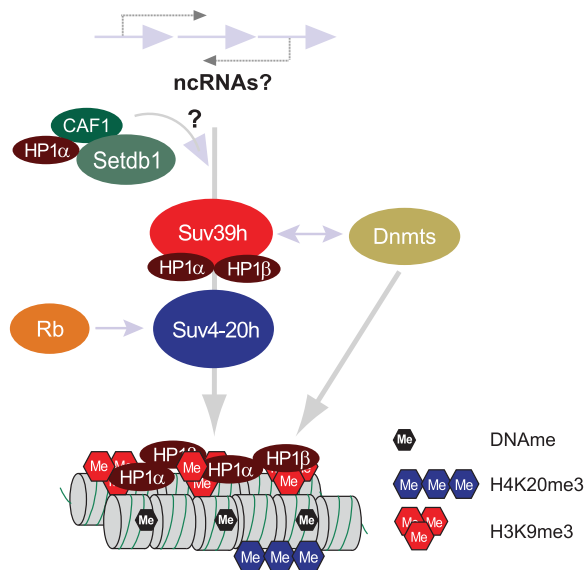


Figure 2. Pericentric heterochromatin marks are established in a sequential pathway (Dambacher et al. 2010)

The Setdb1/CAF1-HP1a complex presumably induces H3K9me1, which is converted to H3K9me3 by Suv39h enzymes. H3K9me3 exerts an effect as a binding platform for HP1 proteins, which in turn recruit Suv4-20h enzymes to induce H4K20me3. Establishment of these modifications is also regulated through interactions with other proteins, for example, members of the retinoblastoma (Rb) family. There is also evidence for interaction of DNA methyltransferases with Suv39h enzymes; however, this interplay needs further characterization.

This leads to recruitment of Clr4, the *S. pombe* homolog of Suv39h, to establish H3K9 methylation at heterochromatin (Grewal and Jia 2007). In mammals, it is still unclear whether a similar link between processing of double-stranded RNA and recruitment of Suv39h exists. H3K9me3 is likely to be established in a highly coordinated manner during replication of pericentric heterochromatin. Data suggest that, in a first step, Setdb1 (SET domain bifurcated 1), in complex with heterochromatin protein 1α (HP1α) and CAF1 (chromatin assembly factor 1), induce H3K9me1 on non-nucleosomal histone H3 (Loyola et al. 2009). *Recently, the methyltransferases Prdm3 and Prdm16 have been shown to also play substantial roles in the generation of H3K9me1 and heterochromatin integrity (Pinheiro et al. 2012).* Subsequently, Suv39h enzymes, which prefer H3K9me1 as substrate, induce H3K9me3, probably even before the H3 molecule is deposited into a nucleosomal context (Rea et al. 2000). Nucleosomal H3K9me3 exerts an effect as a binding platform for HP1 proteins, which in turn recruit Suv4-20h enzymes to establish H4K20me3 (Schotta et al. 2004). The direct interaction of Suv4-20h enzymes with HP1 is necessary to induce H4K20me3; however, interactions with other proteins, for example, members of the retinoblastoma family, can contribute to the establishment of this modification (Gonzalo et al. 2005).

What are the functions of histone lysine methylation marks at pericentric heterochromatin? Although considered highly compact and transcriptionally silent, there is increasing evidence for controlled transcription across pericentric heterochromatin (Eymery et al. 2009). *Transcription factor binding sites for e.g. Pax3 were mapped to major satellite repeats, which could mediate heterochromatin silencing. It was suggested, that distinction between euchromatin and heterochromatin is likely based on a synergistic organization versus a more random distribution of transcription factor binding sites (Bulut-Karslioglu et al. 2012).* Suv39h double-null mutants show enhanced amounts of major satellite transcripts (Martens et al. 2005), indicating that H3K9me3 has an important role in controlling the transcript levels from these pericentric repeat regions. How this control is accomplished, whether H3K9me3 or its binding factors hinder access to RNA Pol II or whether RNA processing or RNA stability are regulated by this modification are some challenging questions in this field. *Interestingly, pericentric H4K20me3 was also described as a late developmental epigenetic marker in mouse embryogenesis (Wongtawan et al. 2011). This finding may imply a potential function of histone lysine methylation at pericentric heterochromatin in developmental transitions.*

1.3 Chromatin changes in development

A detailed review “Epigenetic regulation of development by histone lysine methylation” of which I am a joint first author, can be found in Chapter 2.2.

In mammals, embryonic development starts with the fertilization of the egg, forming the zygote. This develops upon continuous cell divisions into a multicellular preimplantation embryo. From this blastocyst's inner cell mass ESCs can be derived. Two important characteristics of ESCs are: 1) the capacity for self renewal and 2) pluripotency, the ability to give rise to all cell lineages of the three germ layers (Niwa 2007). The chromatin of stem cells is believed to have a more “open” conformation and a higher plasticity, meaning a hyperdynamic chromatin binding of structural proteins like HP1, linker histone H1 or core histones (Meshorer and Misteli 2006; Meshorer et al. 2006). These features are progressively lost upon transition into the multitude of lineage-specific stem cells, progenitors and terminally differentiated cells. Differentiation correlates with an increase in heterochromatin formation and is accompanied by changes in nuclear and chromatin architecture and distribution of heterochromatin proteins (Dialynas et al. 2007; Bhattacharya et al. 2009; Joffe et al. 2010). Lineage-specific gene transcription needs to be activated while lineage-inappropriate genes must be silenced. Chromatin regulation is crucial to control

these differential gene expression patterns. The best studied groups of proteins balancing the developmental regulators are the Trithorax (TrxG) and Polycomb (PcG) systems. TrxG and PcG proteins have been shown to play key roles in developmental gene regulation (Lee et al. 2006; Schuettengruber et al. 2011). Members of these groups can induce, bind and modulate either the active H3K4me3 or the repressive H3K27me3 marks.

1.4 Polycomb gene silencing via H3K27 methylation

The following subchapter and Figure 3 have been published in (Dambacher et al. 2010), a review article, of which I am joint first author (for author contributions see page 8). Slight changes in the text (marked in *italics*) have been conducted.

H3K27 methylation only exists in multicellular organisms and has probably evolved as a system to facilitate cell-type differentiation. Surprisingly, in embryonic stem (ES) cells, H3K27me3 can coexist in the same region with H3K4me3 (Bernstein et al. 2006). Genes that carry this 'bivalent' modification are mainly developmental regulators. Although bivalent genes are repressed, they carry engaged but stalled RNA Pol II. In differentiated cells, bivalent chromatin domains are reduced and genes that are active or repressed are characterized by H3K4me3 or H3K27me3, respectively (Mikkelsen et al. 2007). In ES cells, the major pluripotency genes *Nanog*, *Sox2* and *Oct4*, are highly expressed and marked with H3K4me3. During differentiation, these genes transiently acquire a bivalent state before they become silenced with H3K27me3 (Pan et al. 2007).

We think that H3K27me3 has a major role for developmental transitions (Figure 3). In pluripotent cells, developmental regulators are repressed by the bivalent modifications H3K27me3 and H3K4me3. After lineage decision, these genes are either active and carry H3K4me3 or they are inactive and show enriched signals for H3K27me3. Apparently, in committed progenitor cells there are at least two categories of H3K27me3-repressed genes. One category is lineage-appropriate genes that need to be activated in later stages of differentiation. Full activation of these genes depends on removal of H3K27me3, likely by histone demethylases like UTX and JMJD3 (Agger et al. 2007; Lan et al. 2007) that would then allow binding of transcriptional activators. Lineage-inappropriate genes, the second category, need to

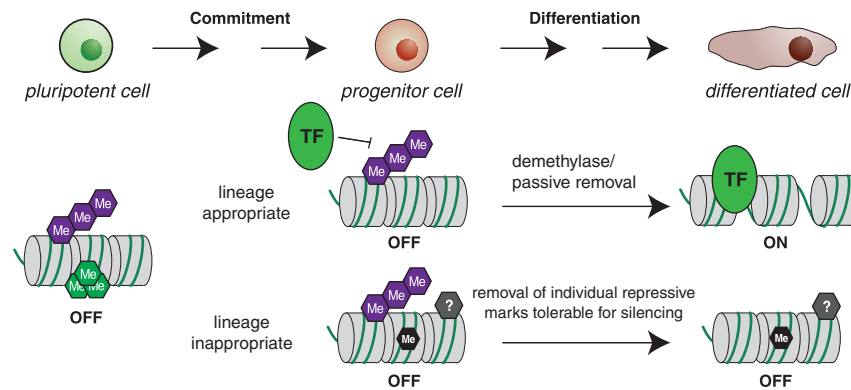


Figure 3. The role of H3K27me3-mediated gene silencing during developmental transitions (Dambacher et al. 2010)

In pluripotent cells, developmental regulators are repressed and carry bivalent H3K4me3 and H3K27me3 modifications. Bivalent marks are reduced in committed progenitor cells in which two categories of H3K27me3-repressed genes exist. Lineage-appropriate genes are activated during terminal differentiation, probably by the active removal of H3K27me3. In contrast, lineage-inappropriate genes are stably repressed by H3K27me3 and other mechanisms, such as different repressive histone modifications and DNA methylation.

be stably repressed by H3K27me3 or together with other mechanisms, like DNA methylation or different repressive histone modifications. Aberrant activation of lineage-inappropriate genes might as well be prevented by the lack of transcriptional activators. *These processes occur within the spatial context of the nucleus. Lately chromosome conformation capture assays provided intriguing insights into large scale in vivo genome organization. These data suggest that the three-dimensional genome positioning in the nucleus affects gene regulation as well (Ferraiuolo et al. 2010).*

1.5 Higher order chromatin structure

Various mechanisms like looping of chromosome territories (Gondor and Ohlsson 2009), chromatin interactions with the nuclear lamina (Burke and Stewart 2012) or molecular crowding (Hancock 2004) can contribute to nuclear architecture. For differential gene expression, the spatial chromatin distribution needs to be dynamically reorganized. Pericentric heterochromatin of different chromosomes clusters to form higher order chromatin structures, called chromocenters, which can be found in most somatic cells. Chromocenters can be visualized as DAPI (4',6-diamidino-2-phenylindole)-dense regions in microscopy (Vissel and Choo 1989; Lehnertz et al. 2003). Initial chromocenter formation in early mouse development seems to be dependent on a strand-specific burst in transcription of the pericentric satellites (Probst et al. 2010). The organization of chromocenters is dynamic and cell

type specific. An increase in chromocenter clustering can be detected upon cellular differentiation (Mayer et al. 2005; Terranova et al. 2005) indicating possible roles in developmental regulation. Consistent with this idea, relocalization of genomic domains to chromocenters can result in gene repression (Brown et al. 1997; Brown et al. 2001) although an explicit silencing mechanism is not known as of yet. Studies have also shown a contribution of the methyl CpG binding protein MeCP2 to global changes in heterochromatin and nucleolar organization during neuronal differentiation (Singleton et al. 2011). A testis-specific histone H1 variant was implicated in chromocenter regulation in spermiogenesis (Catena et al. 2006). Moreover, chromocenter integrity was recently shown to be depending on the histone H3K9-specific methyltransferases Prdm3 and Prdm16 (Pinheiro et al. 2012). However, the roles and biological relevance of chromocenters in genome organization and gene regulation remain elusive.

A dramatic reshaping of the nuclear genome organization occurs in mitosis. An integral part for correct chromosome segregation is an intact pairing of the sister chromatids. This sister chromatid cohesion is mediated by cohesin, a multisubunit complex consisting of a heterodimeric ring of Smc1 and Smc3 proteins and additional stabilizing proteins Scc1 and Scc3 (Onn et al. 2008). In vertebrates, cohesin is loaded in the G1 cell cycle phase and establishes sister chromatid cohesion in S-phase. In fission yeast, cohesin can be targeted to heterochromatic regions by Swi6/HP1 (Nonaka et al. 2002). In mammals, though, there seems to be no interaction between HP1 and cohesin (Koch et al. 2008). However, it was shown recently, that upon knock-down of HP1 proteins sister chromatid cohesion is defective, concomitant with a loss of cohesin proteins at pericentromeres (Shimura et al. 2011). How cohesin is eventually maintained and recruited to mammalian pericentric heterochromatin is yet an open question.

1.6 Centromeres and heterochromatin

Pericentric heterochromatin regions, consisting of major satellite repeats, are localized adjacent to centromeres and have a significant function for proper chromosome segregation and mitotic sister chromatid cohesion (Bernard et al. 2001; Peters et al. 2001; Nonaka et al. 2002; Guenatri et al. 2004). This proximity might be important for correct centromere function as studies in yeast suggest (Olszak et al. 2011), but more detailed analyses are necessary to better understand a possible interplay of these chromosomal domains.

Centromeres are essential chromosome regions that are crucial for correct spindle attachment and chromosome segregation in cell division. In yeast, the incorporation of the centromere-specific histone variant CenH3, is DNA sequence dependent. However this DNA specificity was lost during evolution (Henikoff et al. 2001) and other processes have to maintain the centromere formation process in higher organisms. More and more evidence points to epigenetic factors triggering the establishment of functional centromeres (Black and Cleveland 2011; Perpelescu and Fukagawa 2011). Like pericentric chromatin, centromeric regions consist of arrays of short tandem repeats. These minor satellite repeat regions are bound by centromeric proteins such as the histone variant Cenp-A (Figure 4A).

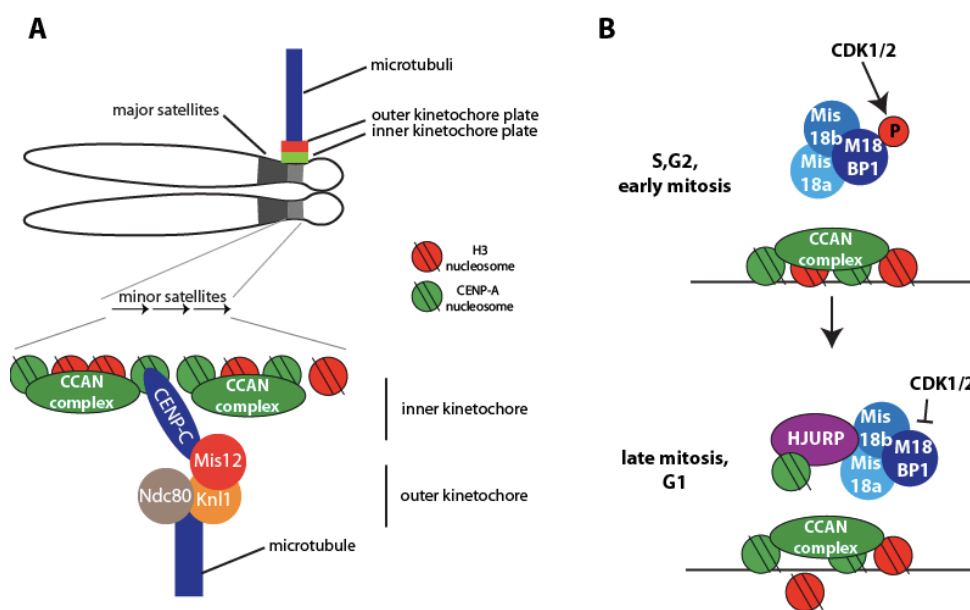


Figure 4. Simplified schemes of centromere composition and CENP-A deposition

A) Chromosomal minor satellite repeats contain Cenp-A nucleosomes. The CCAN protein complex is part of the inner kinetochore region. Cenp-C is linking to the outer kinetochore, which is crucial for microtubule-anchoring in mitosis B) Cenp-A deposition only happens in late mitosis, G1 phase. This coincides with centromere recruitment of the Mis18 complex, which is important for Hjurp loading. Phosphorylation of M18bp1, a member of the Mis18 complex, likely by Cdk1/2, may regulate centromeric Mis18 localization.

Cenp-A incorporation into centromeric chromatin is crucial for centromere and kinetochore function and is cell cycle dependent (Figure 4B). In S-phase, Cenp-A nucleosomes are diluted between sister chromatids. If this distribution is random or if it happens in a semi-conservative way is still unknown. To maintain a functional centromere, Cenp-A levels need to be restored following S-phase, which happens after mitosis in early G1 (Jansen et al. 2007). Cenp-A recruitment is facilitated by various proteins like the Mis18 complex, Mis12 and the histone chaperone Hjurp

(Hayashi et al. 2004; Fujita et al. 2007; Dunleavy et al. 2009; Foltz et al. 2009). It seems that actively transcribed chromatin is fundamental for proper Hjurp function (Bergmann et al. 2011). For Cenp-A incorporation, mediated by the Hjurp chaperone, the Mis18 complex is necessary. This complex consisting of Mis18a, Mis18b and M18bp1, is important for the “priming” of these genomic regions (Fujita et al. 2007) and subsequent kinetochore assembly. It was demonstrated that the localization of the Mis18 complex to centromeres in mitotic telophase to G1 phase is dependent on phosphorylation of M18bp1 (Silva et al. 2012), which may be the crucial regulatory step for the consecutive Cenp-A deposition. Since none of the Mis18 complex members can interact with Cenp-A, the details of this recruitment mechanism to centromeric chromatin are still unclear and may require additional proteins. The constitutive centromere-associated network (CCAN) consists of 16 chromatin proteins, that localize to the centromeric chromatin and may link Cenp-A to the kinetochore (Figure 4A). One of the CCAN members is Cenp-C, a centromere protein which was described to bind DNA and to act as a linker to the outer kinetochore Mis12 complex (Screpanti et al. 2011). It has been shown that Cenp-C and Cenp-T could recruit outer kinetochore proteins to establish an ectopic centromere. But detailed deposition mechanisms for Cenp-A or Cenp-C are still not clear and future experiments have to explain the establishment and maintenance of centromeres and kinetochores.

1.7 Aims of this work

The main objective of this thesis was to get a better understanding of the formation and biological relevance of centromeric and pericentromeric heterochromatin and its protein networks.

Heterochromatin is known to be an important regulator of developmental transitions. However, no mechanistic role of Suv4-20h activity in vertebrate development was known so far. **This study will elaborate on the involvement of Suv4-20h and H4K20me3 in the control of development.** Suv4-20h depletion in *Xenopus laevis* embryogenesis was analyzed and a mouse ES cell to embryoid body differentiation system was established to explore Suv4-20h's impact on germ layer differentiation.

Beyond developmental gene regulation, heterochromatin is also known to be important for genomic integrity and chromatin organization. Pericentromeric heterochromatin is established, maintained and regulated by a complex interplay of proteins. **What are the functions of the heterochromatin protein network and how does it contribute to genomic integrity?** To address these questions, heterochromatin protein dynamics was examined. Furthermore, Suv4-20h knock-out cell lines were analyzed, revealing essential tasks of Suv4-20h in chromosome segregation, nuclear architecture, chromatin compaction and cohesin recruitment.

The influence of heterochromatin on cellular processes is not arising from histone modifying enzymes alone. The histone modifications themselves can affect cellular actions by recruiting a plethora of additional regulatory proteins. **In this dissertation, the novel histone modification H3K56me3 was identified as a conserved pericentromeric heterochromatin mark.** Moreover, the enzymatic framework establishing and removing H3K56me3 could be described.

It is assumed that epigenetic mechanisms "licence" centromeric chromatin for Cenp-A incorporation into chromatin. **The last part of this thesis addresses the role of Cenp-C in recruiting M18bp1 to centromeric chromatin.** Detailed microscopy analyses and protein interaction studies were performed to describe the interplay between Cenp-C and M18bp1.

2. Results

2.1 Heterochromatin dysregulation in human diseases.

Review article: Hahn et al. 2010, Journal of Applied Physiology

HIGHLIGHTED TOPIC | *Epigenetics in Health and Disease*

Heterochromatin dysregulation in human diseases

Matthias Hahn,* Silvia Dambacher,* and Gunnar Schotta

Munich Center for Integrated Protein Science (CiPS^M) and Adolf-Butenandt-Institute, Ludwig-Maximilians-University, Munich, Germany

Submitted 19 January 2010; accepted in final form 29 March 2010

Hahn M, Dambacher S, Schotta G. Heterochromatin dysregulation in human diseases. *J Appl Physiol* 109: 232–242, 2010. First published April 1, 2010; doi:10.1152/jappphysiol.00053.2010.—Heterochromatin is a repressive chromatin state that is characterized by densely packed DNA and low transcriptional activity. Heterochromatin-induced gene silencing is important for mediating developmental transitions, and in addition, it has more global functions in ensuring chromosome segregation and genomic integrity. Here we discuss how altered heterochromatic states can impair normal gene expression patterns, leading to the development of different diseases. Over the last years, therapeutic strategies that aim toward resetting the epigenetic state of dysregulated genes have been tested. However, due to the complexity of epigenetic gene regulation, the “first-generation drugs” that function globally by inhibiting epigenetic machineries might also introduce severe side effects. Thus detailed understanding of how repressive chromatin states are established and maintained at specific loci will be fundamental for the development of more selective epigenetic treatment strategies in the future.

epigenetics; heterochromatin; epigenetic therapy; FSHD; FRDA; cancer

ABNORMAL GENE EXPRESSION patterns are often implicated in the development of different diseases, and thus detailed insights into the underlying mechanisms of gene regulation will be fundamental in identifying novel interventional treatment strategies. Enormous progress over the last few years, particularly in the field of epigenetic regulation, has contributed to a better understanding of how gene activity is controlled.

In the nucleus DNA is packaged with histones and nonhistone proteins in a dynamic structure called chromatin. The grade of chromatin compaction can influence DNA accessibility to the transcriptional machinery. Euchromatin is loosely packed and transcriptionally active, whereas heterochromatin represents a more densely packed chromatin state that is characterized by low transcriptional activity. Heterochromatin can be subdivided into two categories: facultative heterochromatin for silencing of developmental genes and the inactive X chromosome in female mammals, and constitutive heterochromatin, which is formed at pericentromeric regions and telomeres.

Five major epigenetic mechanisms are capable of establishing and stabilizing open or closed chromatin structures, thereby regulating transcriptional activity (Fig. 1).

1) Nucleosome remodelers can stimulate transcription by removing nucleosomes from promoter regions, allowing transcription factors to gain access to the underlying DNA (47).

Additional functions of nucleosome remodelers include histone variant exchange and nucleosome sliding (93).

2) Mammalian cells express three histone H3 variants: H3.1, H3.2, and H3.3. Specific enrichment of H3.3 in transcriptionally active genes and regulatory regions, and H3.1 in repressed or inactive genetic elements suggest a regulatory function of these variants (60). Incorporation of histone variants is also important for many other chromatin-related processes. Faithful DNA damage repair, for example, depends on the presence of a histone H2A variant, H2A.X, which is phosphorylated on damage detection allowing binding of the DNA damage repair machinery (11).

3) Increasing evidence suggests important functions for long noncoding RNAs (ncRNAs) in transcriptional regulation. Direct interactions of chromatin-modifying enzymes with noncoding RNAs are speculated to facilitate targeting to specific genomic loci. A well-studied example is the histone methyltransferase Ezh2, which was shown to interact with different ncRNAs to induce X inactivation and repression of developmental genes (75, 101). Small ncRNAs, e.g., siRNAs and promoter antisense RNAs, can also trigger formation of repressive chromatin structures as will be discussed below.

4) Cytosine bases, preferably in the context of CpG dinucleotides, can be methylated to 5-methylcytosine by DNA methyltransferases. DNA methylation is a repressive modification that is enriched at promoter regions of genes and noncoding DNA sequences. MBD (methyl binding domain) proteins and MeCP2 can bind methylated DNA stretches and in turn recruit corepressor complexes to facilitate transcriptional silencing (78).

* M. Hahn and S. Dambacher contributed equally to this work.

Address for reprint requests and other correspondence: G. Schotta, Adolf-Butenandt-Institute, Ludwig-Maximilians-Univ., Schillerstrasse 44, 80336 Munich, Germany (e-mail: gunnar.schotta@med.uni-muenchen.de).

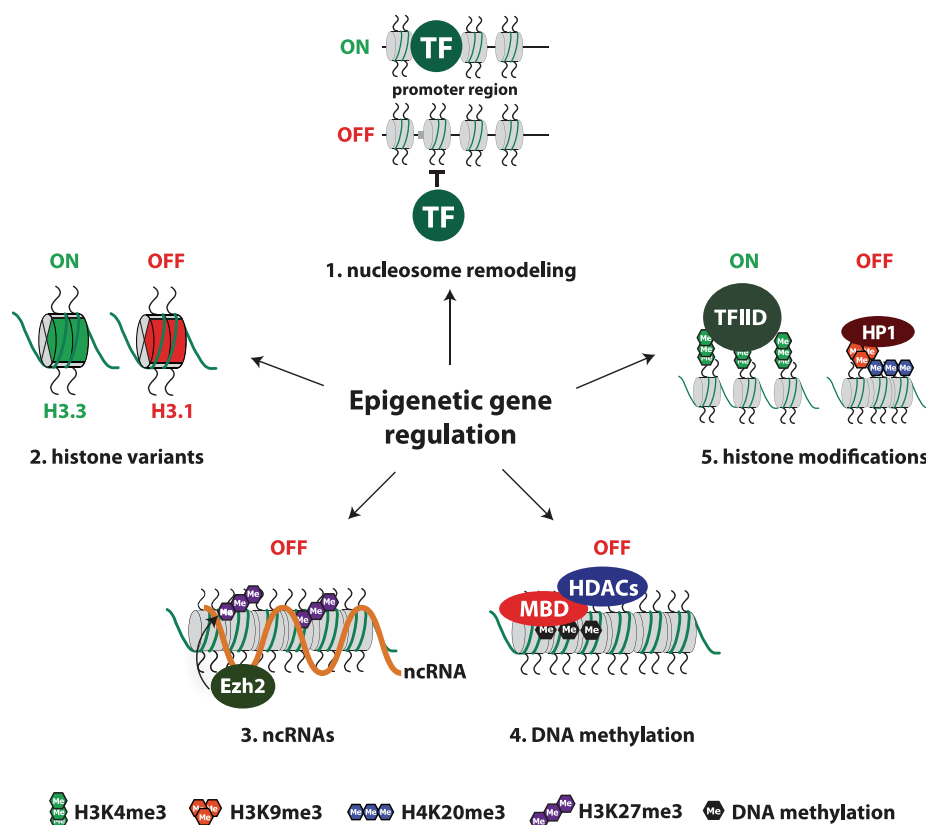


Fig. 1. Major epigenetic mechanisms that regulate chromatin structure. 1) Nucleosomes can prevent binding of transcription factors (TF). One function of nucleosome remodeling complexes is to evict nucleosomes to make promoters accessible. 2) Distinct histone H3 variants associate with transcriptionally active or repressive chromatin domains. 3) Noncoding RNAs (ncRNAs) can recruit histone-modifying enzymes. For example, interaction of the H3K27-specific histone methyltransferase Ezh2 with a ncRNA is important to induce X inactivation. 4) DNA methylation is a repressive modification that is recognized by specific binding factors (MBD proteins) which recruit other corepressor proteins, e.g., HDACs. 5) Histone modifications can recruit binding partners that induce active or repressive chromatin states. H3K4me3 correlates with transcriptional activity and is recognized by components of the TFIID complex. Heterochromatic chromatin structures feature the combinatorial mark H3K9me3 and H4K20me3.

5) Histones can be posttranslationally modified. The major modifications include phosphorylation, acetylation, and methylation. Combinations of different histone modifications represent chromatin signals that are recognized by specific binding proteins that then mediate downstream effects. In the context of transcriptional regulation, histone lysine methylation has been particularly well characterized. This modification generates a high complexity of signals as each lysine position can be mono- (me1), di- (me2), or trimethylated (me3) and distinct binding proteins for each methylation state can mediate different functions. Transcriptionally active, euchromatic domains are characterized by histone H3 lysine 4 trimethylation (H3K4me3) at gene promoters and H3K36me3 across gene bodies (2). The two types of heterochromatin carry distinct modification patterns. Facultative heterochromatin is marked by high levels of H3K27me3 (86). In contrast, constitutive heterochromatin features the combinatorial mark H3K9me3 and H4K20me3 (52).

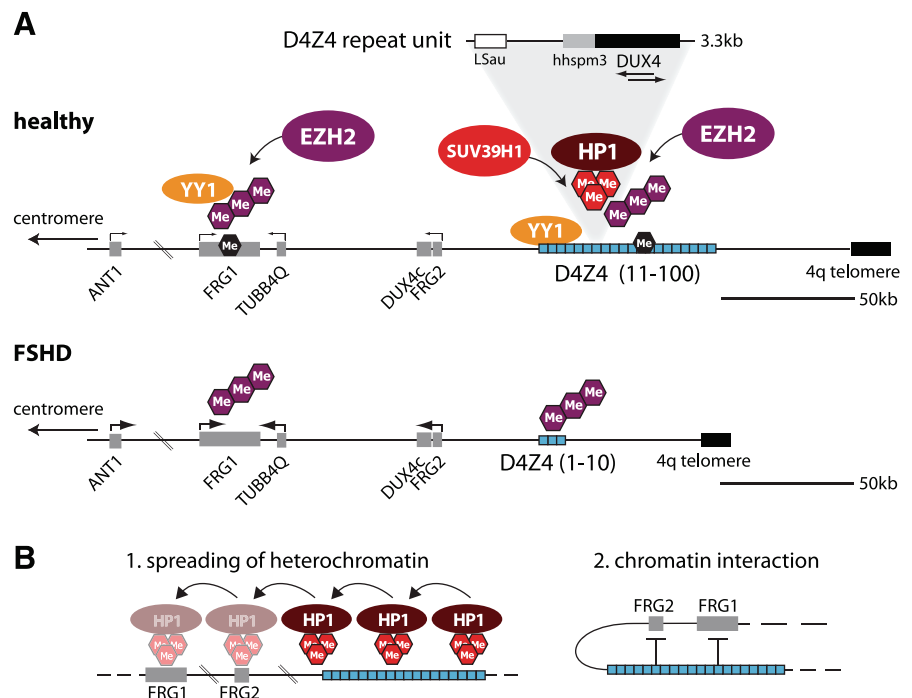
The enzymatic systems for many chromatin-modifying activities have been identified during the last years; however, still very little is known about their targeting to specific genomic regions. For gene silencing and heterochromatin formation across repetitive elements, the production of distinct types of noncoding RNA, antisense RNA and siRNA, have been implicated in the targeting mechanism.

In mammals, the promoter regions of many genes are transcribed at a low level, giving rise to promoter-associated RNAs (30). Antisense ncRNAs that are complementary to these promoter transcripts can mediate recruitment of repressive epigenetic modifications, such as H3K27me3 and H3K9me2. This form of transcriptional gene silencing involves the small RNA binding proteins Argonaute 1 and 2, suggesting that epigenetic machineries interact with components of the RNAi pathway (38, 42).

Targeting of repressive epigenetic modifications to repetitive elements at pericentric heterochromatin appears to involve small interfering RNAs. This mechanism is well understood in the budding yeast, *Schizosaccharomyces pombe*: bidirectional transcription of repetitive elements generates double-stranded RNAs that are cleaved to siRNAs by the endonuclease *dicer*. These small RNAs are then loaded onto *argonaute 1*, which recognizes the DNA sequence from which the small RNA was generated. The histone methyltransferase (HMTase) Clr4 directly interacts with *argonaute 1* and induces repressive H3K9 methylation (29). It is still controversial if this siRNA mechanism exists in mammals; however, there are parallels. An early step in establishing the combinatorial histone modifications at pericentric heterochromatin is H3K9 trimethylation by Suv39h1 and Suv39h2, the mammalian Clr4 homologs. H3K9me3 is recognized by heterochromatin protein 1 (HP1), which, in turn, recruits other HMTases, Suv4-20h1 and Suv4-20h2, to mediate H4K20me3 (81, 82). Heterochromatin in mammals features another important epigenetic modification, DNA methylation. Interestingly, there is interdependence between DNA methylation and H3K9me3 (56). However, the mechanisms are not clear.

An important hallmark of constitutive heterochromatin is its ability to spread from a nucleation site into neighboring regions. Position-effect variegation in *Drosophila* is a well-established model system that highlights this feature (80). In this organism, heterochromatin can eventually expand several hundred kilobases into euchromatic regions. Spreading is hypothesized to work through the following mechanism: H3K9me3 on one nucleosome recruits HP1 which can dimerize and interact with Suv39h that establish H3K9me3 on the neighboring nucleosome. The border between heterochromatic

Fig. 2. Reduced heterochromatin formation in facioscapulohumeral muscular dystrophy (FSHD). A: the subtelomeric region of chromosome arm 4q contains a large array of D4Z4 repeats that consists of repetitive elements (LSau and hhspm3) and the *DUX4* pseudogene. In healthy individuals, 11–100 tandem repeats of D4Z4 induce a local heterochromatic environment that represses transcription of genes across a region of more than 200 kb. DNA methylation, H3K9me3 and HP1, hallmarks of pericentric heterochromatin, can be detected at D4Z4 repeats. There is evidence that DNA specific binding of YY1 recruits EZH2 to induce H3K27me3, another repressive modification. In FSHD patient cells, the reduced D4Z4 array is not able to nucleate a heterochromatic structure. H3K27me3 is still detectable in the locus; however, loss of H3K9me3 and DNA methylation leads to overexpression of genes in the 4q region. B: models for heterochromatin-mediated gene silencing in the 4q region. 1) D4Z4 arrays can induce a local heterochromatic domain that spreads across the 4q locus, thereby inactivating the genes in this region. 2) An alternative explanation is the formation of chromatin interactions between the D4Z4 repeats and gene promoters, also leading to repression.



and euchromatic domains is likely to be dynamically regulated by a balance between the antagonizing activities of activating and repressive factors (19). Heterochromatin spreading also exists in mammals. However, it is not clear to what extent gene regulation is affected by its impairment.

Gene silencing through heterochromatin is an important mechanism to ensure establishment of cell type-specific gene expression patterns. Dysregulation of heterochromatin can result in severe developmental defects. In this review we will discuss human diseases that are connected to defective heterochromatin formation. We will distinguish four different categories: 1) gene overexpression due to reduced heterochromatin, 2) gene silencing through aberrant heterochromatin formation, 3) stable heterochromatin at regulatory elements, and 4) global dysregulation of heterochromatin.

Glossary

AS	Angelman syndrome
DM1	myotonic dystrophy type 1
FRDA	Friedreich's ataxia
FSHD	facioscapulohumeral muscular dystrophy
FXN	frataxin
FXS	fragile X syndrome
HDAC	histone deacetylase
HMTase	histone methyltransferase
HP1	heterochromatin protein 1
IAP	intracisternal type A particle
ICR	imprinting control region
MBD	methyl binding domain
ncRNA	noncoding RNA
PWS	Prader-Willi syndrome
Rb	Retinoblastoma protein
SAHF	senescence-associated heterochromatic foci

REDUCED HETEROCHROMATIN FORMATION IN FACIOSCAPULOHUMERAL MUSCULAR DYSTROPHY

Genes are not only controlled by specific transcriptional activators and repressors. Transcriptional regulation is probably always adapted to the genomic environment of a particular gene. In the first part of this review we will highlight a disease that is caused by reduced formation of heterochromatin across a repetitive region, leading to aberrant activation of genes that are in close proximity.

Facioscapulohumeral muscular dystrophy (FSHD) is a frequent (1:20,000) dominant autosomal disease that is characterized by progressive, often asymmetric weakness and wasting of facial, shoulder, and upper arm muscles (45). FSHD is not caused by mutation of a specific disease gene. Instead, increasing evidence suggests a significant role for a complex epigenetic mechanism, resulting in perturbation of heterochromatic gene silencing in the subtelomeric domain of the long arm of chromosome 4 (24). This region (Fig. 2A) contains several genes (e.g., *FRG1*, *FRG2*, *ANTI*) and pseudogenes (*TUBB4Q*, *DUX4c*) next to a large array of repetitive sequences (D4Z4). Each D4Z4 repeat is ~3.3 kb in length and harbors two classes of the GC-rich repeat sequences hhspm3 and LSau as well as a *DUX4* pseudogene. Furthermore, D4Z4 is overall very GC-rich and has characteristics of CpG islands (35).

Reduced D4Z4 repeat number leads to overexpression of genes in FSHD patients. In healthy individuals up to 100 tandem copies of the D4Z4 element generate a heterochromatic domain that silences the nearby genes (Fig. 2A). FSHD patients have less D4Z4 repeats (1–10 copies) (89). Interestingly, smaller D4Z4 arrays result in earlier disease onset and enhanced severity in patients (74, 85, 97). Genes proximal to D4Z4 are often inappropriately overexpressed in FSHD patients (24). As several of these genes might contribute to the FSHD phenotype, individual overexpression of *FRG1*, *FRG2*,

and *ANT1* was conducted in the mouse model. Notably, overexpression of a single gene, *FRG1*, was sufficient to induce a muscular dystrophy phenotype that mimics human FSHD (23). Loss of function experiments in *Xenopus laevis*, using morpholinos against *xFrg1*, resulted in disrupted organization and inhibited growth of the myotome, and consistent with the mouse data, elevated *xFrg1* levels lead to abnormal muscle formation (31). Not much is known about the molecular functions of FRG1. The protein localizes to nucleoli and Cajal bodies and is probably involved in pre-mRNA splicing (90, 91). Interestingly, *FRG1* overexpression affects splicing of muscle-specific genes (23).

Recent studies have examined more FSHD patients, and the data are very complex. Besides the classic FSHD type with reduced D4Z4 repeats there is also a phenotypic type where repeat length is not altered. Finally, there are also FSHD patients where *FRG1* is rather normally expressed and muscle lineage genes from unrelated loci are dysregulated (39, 46). These studies indicate that the FSHD phenotype can be caused by many different defects; however, in this review we will focus on the mechanism of *FRG1* derepression.

Silencing of FRG1 is dependent on heterochromatic histone methylation marks. The close proximity to D4Z4 repeats and the dependence on the repeat number suggest that *FRG1* expression is affected by epigenetic mechanisms that act in this environment. Not much is known about epigenetic modifications in the 4q35 region (Fig. 2A). In early studies the Polycomb group protein YY1 has been detected at the D4Z4 repeats (24) and at the *FRG1* promoter (5). Recruitment of YY1 correlates with establishment of a repressive histone modification, H3K27me₃, at both regions. In mammals, H3K27 methylation is controlled by the histone methyltransferases EZH1 and EZH2. Although direct evidence is lacking, these enzymes are likely to be responsible for this modification at 4q35. In FSHD patients overexpressing *FRG1*, H3K27me₃ can still be detected at the *FRG1* promoter, suggesting that this modification is not sufficient to induce silencing (5, 100).

Two other repressive epigenetic modifications, DNA methylation and H3K9me₃, might play important roles in this context: DNA methylation has been detected at both D4Z4 repeats and at the *FRG1* promoter, and compared with healthy controls, FSHD patients show significant hypomethylation of these regions (92). Interestingly, hypomethylation is not necessarily dependent on reduced D4Z4 repeats, as patients with normal repeat numbers can also display hypomethylation. In a recent publication Zeng et al. could show that H3K9me₃ is highly enriched at D4Z4 repeats (100), suggesting that this domain has features of constitutive heterochromatin. In line with this model, the authors could demonstrate that, just as in pericentromeric heterochromatin, H3K9me₃ is induced by SUV39H enzymes and can recruit HP1 proteins.

Regulation of FRG1 expression by D4Z4 heterochromatin: spreading vs. locus interactions. How can changes in the D4Z4 repeat number alter epigenetic modifications across the 4q35 region? We hypothesize that the D4Z4 repeat structure induces a heterochromatin environment, using similar mechanisms as found at pericentric heterochromatin. Bidirectional transcription across D4Z4 might generate double-stranded RNAs (14) that can be processed by the RNAi machinery. The resulting small interfering RNAs could recruit SUV39H enzymes to induce H3K9me₃, the binding platform for HP1 proteins.

Spreading of this heterochromatic structure may then establish repression of neighboring genes, such as *FRG1* (Fig. 2B). Reduced D4Z4 repeats might not be able to nucleate enough heterochromatin components to facilitate spreading to very distant loci. This idea is consistent with heterochromatin expansion in *Drosophila* where small heterochromatic regions can spread only around a few kilobases (36); however, special rearrangements that involve large blocks of pericentric heterochromatin can induce spreading across hundreds of kilobases (19). There is also an alternative explanation for D4Z4-dependent silencing of *FRG1*. Chromosome conformation capture (3C) techniques have revealed interactions between the heterochromatic D4Z4 repeats and the *FRG1* promoter (5, 72). These chromatin interactions might induce repression of FRG1 without the necessity of a spreading mechanism (Fig. 2B). If the D4Z4 array is small, interactions with FRG1 might be less frequent and could thus influence the transcriptional status of *FRG1*.

In summary, the data suggest that FSHD is connected to epigenetic dysregulation of D4Z4 repeats; however, at the current stage of analysis it is not possible to build a consistent model. In future studies it will be crucial to analyze a broader spectrum of epigenetic modifications across the entire 4q35 region. Furthermore it would be very helpful to put these data in context with expression analyses of not only *FRG1* but also all other genes in this locus. It is very likely that *FRG1* overexpression only correlates with a subset of FSHD patients and other genes might also contribute to disease development. Before specific disease treatment strategies can be developed, the phenotypic FSHD subtypes need to be much better defined on the molecular level.

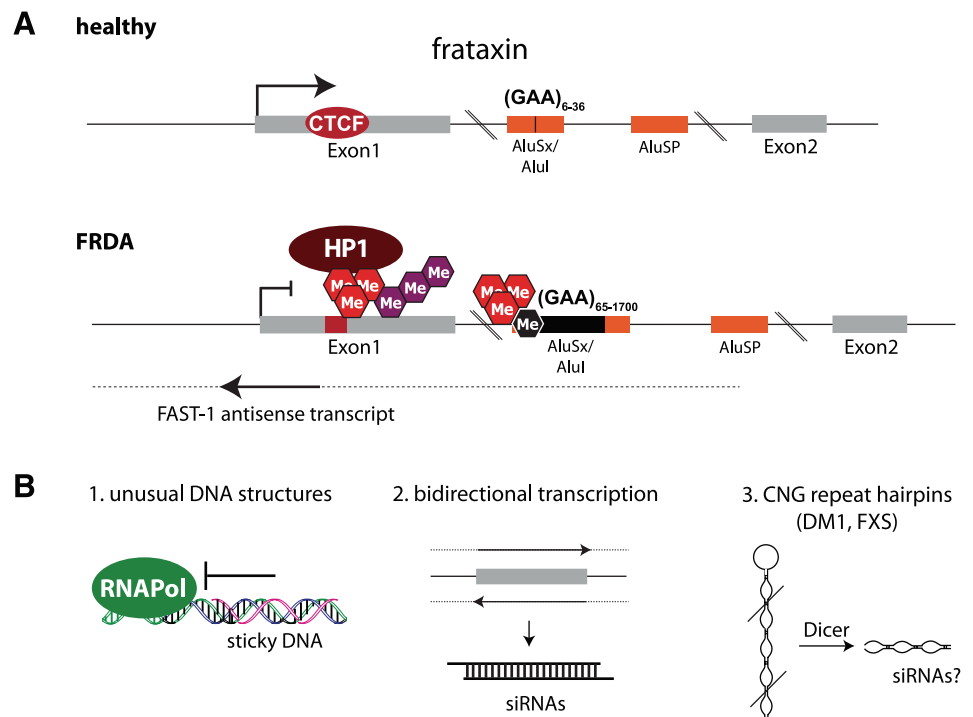
TRIPLET REPEAT INDUCED HETEROCHROMATIN FORMATION IN FRIEDREICH'S ATAXIA

In the previous section we discussed a disease that is caused by reduced heterochromatin silencing. However, there are also human diseases that can be linked to ectopic heterochromatin formation and aberrant gene silencing. In this section we will describe how repetitive expansion of small nucleotide stretches, often triplets, in promoter, intron, or exon regions can lead to heterochromatin formation and silencing of the associated gene, resulting in the development of a disease (59).

Friedreich's ataxia (FRDA) is an autosomal recessive neurodegenerative disease with a frequency of 1:50,000. Patients suffer from progressive ataxia, muscle weakness, heart disease, and eventually diabetes (18, 32). FRDA is caused by an expansion of GAA repeats in intron 1 of frataxin (*FXN*), resulting in substantially reduced transcription of this gene (4, 10). *FXN* is a mitochondrial protein that probably works as an iron chaperone. Its precise function is not entirely clear; however, loss of *FXN* leads to metabolic disturbances, e.g., increased oxidative stress, reduced iron-sulfur clusters, and defects in energy metabolism (67).

GAA repeat expansions cause heterochromatin-induced silencing of frataxin. Healthy individuals carry between 6 and 36 GAA repeats in intron 1 of frataxin (Fig. 3A), while FRDA patients show expansions ranging from 120–1,700 repeats (10, 15, 18, 61). In FRDA patient cells, frataxin is silenced by GAA repeat-dependent formation of heterochromatin (16, 34). Interestingly, repeat length correlates with disease severity (21, 62).

Fig. 3. Triplet repeat induced heterochromatin formation in Friedreich's ataxia (FRDA). **A:** in healthy individuals, intron 1 of the frataxin gene contains only a small number of GAA repeats. The frataxin gene is transcribed, and CTCF binding in exon 1 may be an important regulatory element for this locus. In FRDA patient cells, expanded GAA repeats lead to formation of a heterochromatic domain, coinciding with loss of CTCF binding and silencing of frataxin expression. Interestingly, repression of frataxin correlates with upregulation of the FAST-1 antisense transcript, which may originate from the AluSP sequence in intron 1. **B:** models for initiation of heterochromatin through triplet repeat expansions. 1) Triplet repeats form unusual DNA structures that may result in stalling of the RNA polymerase (RNAPol). This could trigger a damage signal leading to inactivation of the locus. 2) Bidirectional transcription across the frataxin locus may produce double-stranded RNA, triggering siRNA-mediated heterochromatin formation. 3) Transcripts containing large CNG triplet repeats form hairpin structures that can be cleaved by dicer. It is currently unclear if the cleavage products function as siRNAs.



This suggests that increased repeats can more easily or more strongly induce silencing. A hallmark of many repeat expansion diseases is somatic instability. With increasing age of the patient, repeats can become longer, which might even accelerate disease progression. Several lines of evidence suggest that this repeat instability is caused by defects during DNA replication, DNA repair, or recombination (59).

Transcriptional control of FXN in healthy individuals has not been analyzed in detail, yet. It is not known if binding sites for specific transcription factors exist (Fig. 3A). Chromatin immunoprecipitation experiments revealed enrichment of the repressive histone modifications H3K9me3 and H3K27me3 at the FXN promoter and in a region upstream of the GAA repeats (34). Similar to pericentric heterochromatin, there is also enrichment for HP1 and DNA methylation in FRDA patient cells (16, 28).

Exon 1 of the frataxin gene contains a binding site for the chromatin insulator protein CTCF. Interestingly, increased DNA methylation or alterations in the chromatin structure in FRDA correlate with loss of CTCF binding (16). CTCF is speculated to play a role in the regulation of an antisense transcript, FAST-1 (Fig. 3A), which may originate from an AluSP sequence more downstream in intron 1 (50), suggesting that antisense transcription is involved in disease development. In agreement with these data FAST-1 cannot be detected in healthy individuals; however, as only a small region of this transcript is known (Fig. 3A, black arrow), it is formally possible that in healthy individuals a shorter form of FAST-1 is transcribed.

Although production of the noncoding FAST-1 RNA at the frataxin locus might influence silencing, heterochromatin formation depends mainly on the GAA repeat expansion. This has been demonstrated in a transgenic mouse model where short GAA repeat expansions induce silencing of a linked marker gene (79). Repression of this transgene correlates with de-

creased promoter accessibility and can be enhanced by HP1 overexpression, typical features of a heterochromatic state. Importantly, silencing of the GAA repeat transgene did not depend on the genomic integration site, suggesting that the heterochromatic state is directly induced by the GAA repeat region.

How can repeat expansions induce heterochromatin? Different models are discussed to explain how small repeat expansions can trigger heterochromatin formation (Fig. 3B); however, there is no ultimate proof for any of them. An interesting feature of long GAA triplet repeats is the formation of non-B-DNA structures and sticky DNA in vitro (94). If such structures, e.g., a DNA triplex, form at the FXN locus, normal transcription could be inhibited and RNA polymerase might stall or progress at a much slower rate. In this case cellular checkpoint mechanisms might trigger heterochromatin formation.

Another intriguing possibility to explain heterochromatin formation is antisense transcription across the FXN locus (Fig. 3B). In *S. pombe* bidirectional transcription generates double-stranded RNA, which is processed by the RNAi machinery and can induce heterochromatin formation across the underlying genomic locus (29). Thus production of the FAST-1 transcript could be crucial for the silencing mechanism in FRDA. Furthermore there is evidence for antisense transcription in other repeat expansion diseases (13, 53), indicating that this process may be of general importance. A major regulator for antisense transcript production might be CTCF. In FRDA patients, loss of CTCF binding correlates with accumulation of FAST-1 transcripts, and also in a related disease, myotonic dystrophy type 1 (DM1), CTCF limits antisense transcription (13). Further analyses are required to demonstrate whether CTCF directly functions in transcriptional control at these loci.

A third model for induction of heterochromatin at trinucleotide repeat loci was suggested recently (48) and might play a

role in distinct repeat expansion diseases (Fig. 3B). Transcripts containing expansions of CAG or CUG repeats, as found in fragile X syndrome (FXS) and DM1, can form stable hairpin loops that are substrates for dicer (48). Reduced transcript levels can be explained by dicer-mediated cleavage of the primary transcripts, leading to lower production of mature RNA. In analogy to siRNA triggered formation of heterochromatin, cleavage products of CNG hairpins might resemble siRNAs that potentially trigger heterochromatin formation using a similar mechanism.

Therapeutic strategies for repeat expansion diseases have focused on the development of treatments to minimize the adverse effects of missing an important gene. Mitochondrial dysfunction in FRDA, for example, benefits from an antioxidant therapy. Over the last years it has become clear that frataxin is silenced by local heterochromatin. This has led, more recently, to novel therapeutic strategies being developed that impair heterochromatin formation to allow higher expression of frataxin. This class of therapeutics, histone deacetylase inhibitors, is very promising as frataxin levels could be increased in FRDA patient cells (34) and in a mouse model (73). However, global dysregulation of histone deacetylation might cause severe side effects. Only a better molecular understanding of the epigenetic machinery that specifically targets repeat expansion loci will help to ultimately develop specific compounds for treatment of this class of diseases.

EPIGENETIC DYSREGULATION OF IMPRINTING CONTROL REGIONS IN PRADER-WILLI SYNDROME AND ANGELMAN SYNDROME

Heterochromatin at regulatory elements can affect transcription of many genes in a large genomic domain. In this section we discuss diseases that are caused by genetic or epigenetic dysregulation at such regulatory regions.

Genomic imprinting is a mechanism to establish allele-specific differences in gene expression dependent on the parental origin (37). Imprinted genes are only expressed from either the maternal or the paternal chromosome. Regulatory elements, so-called imprinting control regions (ICRs), display parental-specific epigenetic profiles. An ICR on one allele can carry active, euchromatic modifications, whereas the ICR of the other allele shows features of heterochromatin. The different epigenetic profiles of maternal and paternal ICRs consequently affect transcription of multiple neighboring genes in *cis* by mechanisms that are not completely understood. Notably, epigenetic programming at ICRs is mitotically very stable (17). Genetic or epigenetic alterations that lead to impairment of an ICR on one allele can result in dysregulation of many genes on the same chromosome, leading to severe developmental defects and disease.

Two imprinting control regions regulate multiple genes in the 15q11–13 region. In humans, two distinct diseases, Prader-Willi syndrome (PWS) and Angelman syndrome (AS), are both caused by aberrant imprinting of genes in the 15q11–13 chromosomal region. Although the same chromosomal region is linked with both diseases, the phenotypic appearance of PWS and AS patients dramatically differs. PWS occurs in 1:15,000–30,000 individuals. Patients display infantile hypotonia, obesity, short stature, small hands and feet, growth hormone deficiency, mental retardation, and behavioral prob-

lems (9). About 70% of PWS patients carry a deletion of the paternal 15q11–13 region; about 30% have a maternal disomy of chromosome 15. The complex phenotypic appearance of PWS is the result of dysregulation of several genes in 15q11–13 (55). AS patients, in contrast, are characterized by microcephaly, ataxia, mental retardation, jerky arm movements, absence of speech, and sleep disorders (95). Similar to PWS about 70% of AS individuals carry 15q11–13 deletions, which always have a maternal origin. Around 10% of AS patients carry mutations in the ubiquitin protein ligase *UBE3A*, suggesting that dysfunction of this single gene can cause AS (44, 57). Small percentages of AS patients have imprinting defects or parental disomy of chromosome 15.

Deletion mapping of AS and PWS patients revealed two imprinting centers in 15q11–13 that are only 35 kb apart (8). The PWS-ICR controls numerous genes in this locus, e.g., *MKRN3*, *MAGEL2*, *NDN*, *SNURF-SNRPN* and a stretch of *snoRNA* genes, whereas the AS-ICR mainly regulates transcription of two very distant genes, *UBE3A* and *ATPC10* (Fig. 4). Genes under control of the PWS-ICR are expressed only from the paternal chromosome. In contrast, *UBE3A* and *ATPC10* which are regulated by the AS-ICR show only maternal expression.

Local heterochromatin formation at the maternal PWS-ICR. How are these allele-specific expression patterns established? The AS-ICR seems central to induce silencing of the maternal PWS-ICR during development (40, 69); however, it might not be involved in direct transcriptional control of the 15q11–13 region, which is rather regulated by the PWS-ICR. On the paternal allele, PWS-ICR carries active epigenetic modifications, i.e., H3K4 methylation (98) and absence of DNA methylation (20, 25). This stimulates transcription in *cis* across the entire 15q11–13 region. Notably, in this case *UBE3A* is transcribed in the antisense direction, which may inhibit the production of sense transcript, and is hypothesized to keep this gene silent on the paternal chromosome (76). On the maternal allele, however, the PWS-ICR features heterochromatic modifications, such as H3K9me3, H4K20me3, and DNA methylation (20, 25, 96, 98). Genes in the 15q11–13 region are consequently silenced by DNA methylation in their promoter regions. The *UBE3A* gene can be transcribed from the maternal allele as there is no more antisense transcription (Fig. 4).

It is not very well known how these epigenetic modifications are established at imprinting control regions. Most data have been obtained from analyses of mouse models; however, it seems plausible that the mechanisms are also conserved in humans. H3K9me3 at the PWS-ICR depends on the histone methyltransferase G9a and, as G9a can only induce H3K9me1 and H3K9me2, another H3K9-specific HMTase is required for the full establishment of this modification (99). H4K20me3 at ICRs is induced by Suv4–20h enzymes using a similar pathway as found in pericentric heterochromatin (68). DNA methylation is established independently of the histone modifications by DNA methyltransferases. The targeting of these enzymes to ICRs is still unclear.

DNA methylation has an important dual role at imprinting centers. It can attract DNA methylation binding proteins, such as MeCP2, and it can also prevent binding of important transcriptional regulators, e.g., CTCF (37). It is still largely unclear how the chromatin state of a relatively small imprinting control region can induce silencing or transcriptional activation

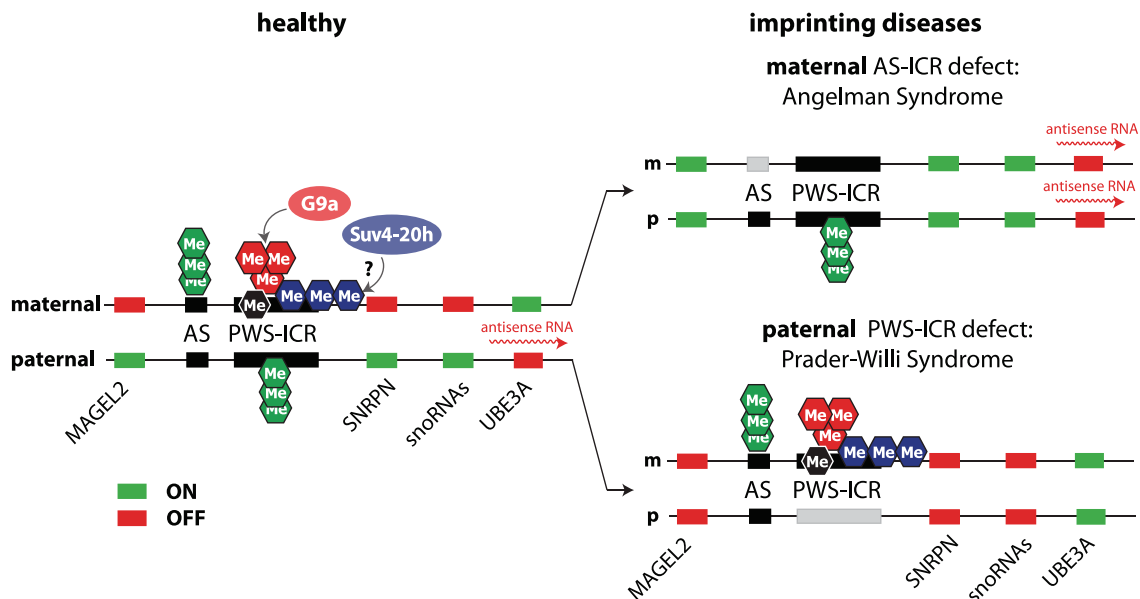


Fig. 4. Prader-Willi syndrome (PWS) and Angelman syndrome (AS) are caused by epigenetic dysregulation of imprinting control regions in chromosomal region 15q11–q13. The 15q11–13 region (not to scale and not all genes are shown) contains 2 imprinting control regions, separated by ~35 kb. On the maternal chromosome, establishment of H3K4me3 at the AS-imprinting control region (ICR) triggers heterochromatic modifications at the PWS-ICR leading to repression of most genes in the locus; only *UBE3A* is maternally expressed. On the paternal chromosome, active modifications at the PWS-ICR stimulate transcriptional activity of 15q11–13 genes. Antisense transcription across *UBE3A* prevents production of the normal sense transcript, thereby silencing the paternal *UBE3A*. In AS patients, inactivation or deletion of the maternal AS-ICR prevents establishment of repressive modification at PWS-ICR. Consequently, 15q11–13 genes are transcribed and production of the *UBE3A* antisense transcript leads to complete inactivation of this gene on both alleles. Deletion or inactivation of the paternal PWS-ICR in PWS patients prevents transcriptional activation of 15q11–13 genes.

of multiple neighboring genes. One possibility is a spreading mechanism of the active or inactive state. Alternatively, formation of chromatin interactions that bring promoter regions of the neighboring genes in close proximity to the imprinting control region might help to transmit epigenetic states onto these target genes.

Stability of allele-specific expression patterns in imprinting diseases. How can defective imprinting on only one allele generate severe diseases as, e.g., AS and PWS? In AS patients, if the AS-ICR is maternally impaired repressive heterochromatin cannot be established on the PWS-ICR (Fig. 4). This leads to aberrant activation of the PWS-ICR on the maternal chromosome, which prevents proper sense transcription of *UBE3A* (76). In PWS patients, deletion or impairment of the paternal PWS-ICR, which would normally carry active modifications, prevents transcriptional activity across the entire 15q11–13 region on the paternal chromosome (55). Notably, all these genes are potentially functional on the maternal chromosome; however, they cannot be expressed due to repressive histone modifications at the maternal PWS-ICR (Fig. 4). An intriguing possibility to treat this imprinting disease would be to reactivate these silenced alleles. Apparently, specific DNA elements in the PWS/AS ICR are necessary to maintain the allelic imprint by recruiting regulatory proteins that are involved in targeting epigenetic machineries (41). A detailed mechanistic knowledge of these mechanisms is required to develop specific treatment strategies. However, a balanced treatment with general inhibitors of DNA methylation (77) or histone methylation (49) might already be able to improve the situation.

HETEROCHROMATIN DYSFUNCTION INCREASES CANCER SUSCEPTIBILITY

Heterochromatin has important functions in gene silencing and genomic stability, two processes that are often impaired in cancer cells. Dysregulation of heterochromatin is expected to increase cancer susceptibility by two mechanisms that we will discuss below: aberrant silencing of tumor-suppressor genes and increased genomic instability by reduced formation of pericentric heterochromatin.

Heterochromatin-induced silencing of tumor-suppressor genes in cancer cells. Numerous reports have provided evidence for aberrant heterochromatin formation at tumor-suppressor genes leading to their silencing (12, 83). Stochastic establishment of epigenetic silencing at such genes could induce mitotically heritable repression, potentially increasing cancer susceptibility. Recently, an intriguing possibility of how such aberrant silencing may be established has been proposed: during DNA damage repair, repressive modifications accumulate at the break site, and in a minor proportion of events, these modifications are not completely removed after successful repair (66).

A well-studied repressive epigenetic modification in the context of tumorigenesis is DNA methylation, which is often dysregulated in cancer cells and connected to aberrant silencing of tumor-suppressor genes (3, 12). Other repressive modifications, e.g., histone lysine methylation, can also be dysregulated in tumor cells. There is evidence that epigenetic silencing of tumor-suppressor genes can be mediated by distinct combinations of H3K27me3, H3K9me2 (58), H3K9me3 (54), and H4K20me3 (51). Reactivation of these aberrantly silenced genes might support a conventional tumor therapy. Inhibition of

DNA methyltransferases (5-aza-cytidine) and histone deacetylases (trichostatin A) can indeed be used to remove these repressive modifications and to consequently reactivate the silenced genes. A very general methylation inhibitor, 3-deazaneplanocin A (DZnep), was shown to be effective in removing H3K27me3 (84). SUV39H1-mediated H3K9me3 in combination with DNA methylation also plays important roles in aberrant silencing of tumor suppressor genes (54). Intriguingly, removal of H3K9me3 by RNAi knockdown of SUV39H1 leads to reactivation of these genes even though DNA methylation was still present at the promoter regions (54). In summary, the data suggest that DNA methyltransferases and histone methyltransferases could be interesting therapeutic targets for cancer treatment.

Reduced formation of pericentric heterochromatin is connected with genomic instability. Heterochromatin has important functions in gene regulation, chromosome segregation, and maintenance of genomic stability. Thus global dysregulation of heterochromatin might also increase cancer susceptibility. In agreement with this hypothesis, a strongly reduced level of H4K20me3 was found to be a hallmark of many human cancers (22, 87, 88), and reduced levels of HP1 α have been found in highly invasive and metastatic breast cancer cell lines (43). It is not clear if these findings represent cause or consequence of tumorigenesis; however, several lines of evidence suggest that impairment of heterochromatin can increase cancer susceptibility.

Cause-effect relationships can be best analyzed in the mouse. Pericentric heterochromatin is impaired in mice lacking the two Suv39h HMTases (*Suv39h* dn mice). Interestingly, these mice are highly susceptible to the development of B cell lymphomas (70). The actual cause of tumorigenesis is not clear; however, four heterochromatin-mediated mechanisms might account for this phenotype.

1) Formation of a stable heterochromatic structure around centromere regions is important to ensure proper chromosome segregation. *Suv39h* deficient fibroblasts display increased aneuploidy, and *Suv39h* dn mice develop lymphomas, which sometimes have a hypertetraploid karyotype, characterized by nonsegregated chromosomes that remain attached through their acrocentric regions (70). An intriguing hypothesis of how aneuploidy accelerates tumorigenesis is that it increases the loss of heterozygosity of tumor suppressor genes (1).

2) Fully established pericentric heterochromatin might be important for the proper repair of DNA damage in satellite regions. Interestingly, DNA damage repair is delayed in *Suv39h* dn mutants, suggesting that impaired heterochromatin is less efficiently repaired (27). It is also possible that DNA damage repair is faulty in these mutants, which could generate mutations that support tumorigenesis.

3) Cells in which DNA damage cannot be fully repaired enter senescence, a stable cell cycle arrest that limits the proliferation of damaged cells and therefore functions as a natural barrier for cancer progression. Senescence is connected to establishment of senescence-associated heterochromatic foci (SAHF), which stably inactivate E2F target genes, thereby preventing cell cycle progression. SAHFs have features of pericentric heterochromatin with enrichment of H3K9me3, H4K20me3, and HP1 (65). Notably, establishment of the senescence program is impaired in absence of *Suv39h1*, leading to increased tumorigenesis (7).

4) Silencing of mobile elements is another mechanism to ensure genome stability. Strong activation of the retrovirus-like intracisternal type A particle (IAP) elements, for example, results in catastrophic defects during meiosis (6). A lower level of activation might lead to integration of transposable elements in tumorigenesis-related genes and could activate oncogenes or mutate tumor suppressor genes. Heterochromatic histone modifications H3K9me3 and H4K20me3 play important roles in silencing LINE-1 and probably other transposable elements. These repressive modifications are mediated by retinoblastoma (Rb) family members that can interact with histone methyltransferases (26). Consequently, dysregulation of Rb function removes H3K9me3 and H4K20me3 from LINE-1 elements, resulting in their activation (63). It remains to be tested if genomic instability through activated mobile elements is a major cause of tumorigenesis.

In summary, these connections, which relate to defective heterochromatin formation, underline the complexity of the tumorigenesis process. Cancer cells always carry genetic and epigenetic mutations, and it seems plausible that epigenetic dysregulation sensitizes cells to accumulate oncogenic genetic mutations. Epigenetic treatment strategies with general histone or DNA methylation inhibitors, which aim for reexpression of tumor suppressor genes, can be very dangerous. Those drugs might kill tumor cells; however, by the same time they sensitize healthy cells to develop a new tumor.

Perspectives

Enormous progress over the last years has improved our understanding of how epigenetic mechanisms contribute to the development of human diseases. These efforts already led to the development of novel therapeutic strategies, which, however, may have side effects. To refine these strategies, we need to better understand the interplay between the epigenetic machineries, and we need to learn more about selective targeting mechanisms. Regarding heterochromatin-mediated gene silencing, there is evidence from *Drosophila* that different chromosomal regions use distinct combinations of epigenetic machineries (71). We are only beginning to understand this interplay in mammals, where a higher complexity of enzymatic systems induces the different epigenetic modifications. Notably, in the context of human diseases, this high complexity suggests that disease-critical pathways may exist, whose inhibition or stimulation would not globally impair epigenetic programming of healthy cells. To identify such pathways, a consistent and systematic analysis of epigenetic modifications and transcriptional regulation in healthy vs. patient material is required.

Identification of epigenetic machineries that work in the context of a specific disease-related locus is only the first step. The next question is how these machineries are targeted. Do these epigenetic enzymes interact with proteins that recognize a specific DNA sequence and could an inhibitor for this interaction prevent targeting? Recent data suggest that a different mechanism, binding of noncoding RNAs, assists in targeting. If we understand how the production of these RNAs is controlled we might get novel tools for regulating site-specific targeting of epigenetic modifications (33, 64). Reagents that mediate degradation of these ncRNAs, e.g., antisense RNAs, might allow specific reactivation of aberrantly

silenced genes, such as frataxin in FRDA or tumor suppressor genes in cancer.

In this review we discuss how heterochromatin dysregulation can generate severe disease phenotypes. Much less is known about smaller changes in epigenetic programming that would not necessarily cause pathological effects. We hypothesize that during aging, stochastic errors in establishing or maintaining epigenetic programming occur during DNA damage repair (66), or can be triggered by environmental factors. Accumulation of this epigenetic damage may then contribute to the systemic aging phenotype. Treatment strategies that aim toward correcting epigenetic programming in human diseases may also be useful to treat the adverse effects of aging.

ACKNOWLEDGMENTS

We apologize to all authors whose work we could not cite due to space limitations. We are grateful to Antonia Jack for critical reading of the manuscript.

GRANTS

Research in the laboratory of G. Schotta is financed by the SFB-TR5 Chromatin, SFB 684, Center of Integrated Protein Science Munich and Friedrich Baur-Stiftung.

DISCLOSURES

No conflicts of interest (financial or otherwise) are declared by the authors.

REFERENCES

- Baker DJ, Jin F, Jeganathan KB, van Deursen JM. Whole chromosome instability caused by Bub1 insufficiency drives tumorigenesis through tumor suppressor gene loss of heterozygosity. *Cancer Cell* 16: 475–486, 2009.
- Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. High-resolution profiling of histone methylations in the human genome. *Cell* 129: 823–837, 2007.
- Baylin SB. DNA methylation and gene silencing in cancer. *Nat Clin Pract Oncol* 2, Suppl 1: S4–S11, 2005.
- Bidichandani SI, Ashizawa T, Patel PI. The GAA triplet-repeat expansion in Friedreich ataxia interferes with transcription and may be associated with an unusual DNA structure. *Am J Hum Genet* 62: 111–121, 1998.
- Bodega B, Ramirez GD, Grasser F, Cheli S, Brunelli S, Mora M, Meneveri R, Marozzi A, Mueller S, Battaglioli E, Ginelli E. Remodeling of the chromatin structure of the facioscapulohumeral muscular dystrophy (FSHD) locus and upregulation of FSHD-related gene 1 (FRG1) expression during human myogenic differentiation. *BMC Biol* 7: 41, 2009.
- Bourc'his D, Bestor TH. Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature* 431: 96–99, 2004.
- Braig M, Lee S, Loddikenemper C, Rudolph C, Peters AH, Schlegelberger B, Stein H, Dorken B, Jenuwein T, Schmitt CA. Oncogene-induced senescence as an initial barrier in lymphoma development. *Nature* 436: 660–665, 2005.
- Buiting K, Lich C, Cottrell S, Barnicoat A, Horsthemke B. A 5-kb imprinting center deletion in a family with Angelman syndrome reduces the shortest region of deletion overlap to 880 bp. *Hum Genet* 105: 665–666, 1999.
- Butler MG. Genomic imprinting disorders in humans: a mini-review. *J Assist Reprod Genet* 26: 477–486, 2009.
- Campuzano V, Montermini L, Molto MD, Pianese L, Cossee M, Cavalcanti F, Monros E, Rodius F, Duclos F, Monticelli A, Zara F, Canizares J, Koutnikova H, Bidichandani SI, Gellera C, Brice A, Trouillas P, De Michele G, Filla A, De Frutos R, Palau F, Patel PI, Di Donato S, Mandel JL, Coccoza S, Koenig M, Pandolfo M. Friedreich's ataxia: autosomal recessive disease caused by an intronic GAA triplet repeat expansion. *Science* 271: 1423–1427, 1996.
- Celeste A, Petersen S, Romanienko PJ, Fernandez-Capetillo O, Chen HT, Sedelnikova OA, Reina-San-Martin B, Coppola V, Meffre E, Difilippantonio MJ, Redon C, Pilch DR, Olaru A, Eckhaus M, Camerini-Otero RD, Tessarollo L, Livak F, Manova K, Bonner WM, Nussenzweig MC, Nussenzweig A. Genomic instability in mice lacking histone H2AX. *Science* 296: 922–927, 2002.
- Chen J, Odenike O, Rowley JD. Leukaemogenesis: more than mutant genes. *Nat Rev Cancer* 10: 23–36, 2010.
- Cho DH, Thienes CP, Mahoney SE, Analau E, Filippova GN, Tapscott SJ. Antisense transcription and heterochromatin at the DM1 CTG repeats are constrained by CTCF. *Mol Cell* 20: 483–489, 2005.
- Clapp J, Mitchell LM, Bolland DJ, Fantes J, Corcoran AE, Scotting PJ, Armour JA, Hewitt JE. Evolutionary conservation of a coding function for D4Z4, the tandem DNA repeat mutated in facioscapulohumeral muscular dystrophy. *Am J Hum Genet* 81: 264–279, 2007.
- Cossee M, Schmitt M, Campuzano V, Reutenauer L, Moutou C, Mandel JL, Koenig M. Evolution of the Friedreich's ataxia trinucleotide repeat expansion: founder effect and premutations. *Proc Natl Acad Sci USA* 94: 7452–7457, 1997.
- De Biase I, Chutake YK, Rindler PM, Bidichandani SI. Epigenetic silencing in Friedreich ataxia is associated with depletion of CTCF (CCCTC-binding factor) and antisense transcription. *PLoS One* 4: e7914, 2009.
- Delaval K, Feil R. Epigenetic regulation of mammalian genomic imprinting. *Curr Opin Genet Dev* 14: 188–195, 2004.
- Durr A, Cossee M, Agid Y, Campuzano V, Mignard C, Penet C, Mandel JL, Brice A, Koenig M. Clinical and genetic abnormalities in patients with Friedreich's ataxia. *N Engl J Med* 335: 1169–1175, 1996.
- Ebert A, Schotta G, Lein S, Kubicek S, Krauss V, Jenuwein T, Reuter G. Su(var) genes regulate the balance between euchromatin and heterochromatin in *Drosophila*. *Genes Dev* 18: 2973–2983, 2004.
- El-Maarri O, Buiting K, Peery EG, Kroisel PM, Balaban B, Wagner K, Urman B, Heyd J, Lich C, Brannan CI, Walter J, Horsthemke B. Maternal methylation imprints on human chromosome 15 are established during or after fertilization. *Nat Genet* 27: 341–344, 2001.
- Filla A, De Michele G, Cavalcanti F, Pianese L, Monticelli A, Campanella G, Coccoza S. The relationship between trinucleotide (GAA) repeat length and clinical features in Friedreich ataxia. *Am J Hum Genet* 59: 554–560, 1996.
- Fraga MF, Ballestar E, Villar-Garea A, Boix-Chornet M, Espada J, Schotta G, Bonaldi T, Haydon C, Ropero S, Petrie K, Iyer NG, Perez-Rosado A, Calvo E, Lopez JA, Cano A, Calasanz MJ, Colomer D, Piris MA, Ahn N, Imhof A, Caldas C, Jenuwein T, Esteller M. Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nat Genet* 37: 391–400, 2005.
- Gabellini D, D'Antona G, Moggio M, Prella A, Zecca C, Adami R, Angeletti B, Ciscato P, Pellegrino MA, Bottinelli R, Green MR, Tupler R. Facioscapulohumeral muscular dystrophy in mice overexpressing FRG1. *Nature* 439: 973–977, 2006.
- Gabellini D, Green MR, Tupler R. Inappropriate gene activation in FSHD: a repressor complex binds a chromosomal repeat deleted in dystrophic muscle. *Cell* 110: 339–348, 2002.
- Glenn CC, Saitoh S, Jong MT, Filbrandt MM, Surti U, Driscoll DJ, Nicholls RD. Gene structure, DNA methylation, and imprinted expression of the human SNRPN gene. *Am J Hum Genet* 58: 335–346, 1996.
- Gonzalo S, Garcia-Cao M, Fraga MF, Schotta G, Peters AH, Cotter SE, Eguia R, Dean DC, Esteller M, Jenuwein T, Blasco MA. Role of the RB1 family in stabilizing histone methylation at constitutive heterochromatin. *Nat Cell Biol* 7: 420–428, 2005.
- Goodarzi AA, Noon AT, Deckbar D, Ziv Y, Shiloh Y, Lobrich M, Jeggo PA. ATM signaling facilitates repair of DNA double-strand breaks associated with heterochromatin. *Mol Cell* 31: 167–177, 2008.
- Greene E, Mahishi L, Entezam A, Kumari D, Usdin K. Repeat-induced epigenetic changes in intron 1 of the frataxin gene and its consequences in Friedreich ataxia. *Nucleic Acids Res* 35: 3383–3390, 2007.
- Grewal SI, Elgin SC. Transcription and RNA interference in the formation of heterochromatin. *Nature* 447: 399–406, 2007.
- Han J, Kim D, Morris KV. Promoter-associated RNA is required for RNA-directed transcriptional gene silencing in human cells. *Proc Natl Acad Sci USA* 104: 12422–12427, 2007.
- Hanel ML, Wuebbles RD, Jones PL. Muscular dystrophy candidate gene FRG1 is critical for muscle development. *Dev Dyn* 238: 1502–1512, 2009.
- Harding AE. Friedreich's ataxia: a clinical and genetic study of 90 families with an analysis of early diagnostic criteria and intrafamilial clustering of clinical features. *Brain* 104: 589–620, 1981.

33. Hawkins PG, Santoso S, Adams C, Anest V, Morris KV. Promoter targeted small RNAs induce long-term transcriptional gene silencing in human cells. *Nucleic Acids Res* 37: 2984–2995, 2009.
34. Herman D, Jenssen K, Burnett R, Soragni E, Perlman SL, Gottesfeld JM. Histone deacetylase inhibitors reverse gene silencing in Friedreich's ataxia. *Nat Chem Biol* 2: 551–558, 2006.
35. Hewitt JE, Lyle R, Clark LN, Valleley EM, Wright TJ, Wijmenga C, van Deutekom JC, Francis F, Sharpe PT, Hofker M, et al. Analysis of the tandem repeat locus D4Z4 associated with facioscapulohumeral muscular dystrophy. *Hum Mol Genet* 3: 1287–1295, 1994.
36. Hines KA, Cryderman DE, Flannery KM, Yang H, Vitalini MW, Hazelrigg T, Mizzen CA, Wallrath LL. Domains of heterochromatin protein 1 required for *Drosophila melanogaster* heterochromatin spreading. *Genetics* 182: 967–977, 2009.
37. Ideraabullah FY, Vigneau S, Bartolomei MS. Genomic imprinting mechanisms in mammals. *Mutat Res* 647: 77–85, 2008.
38. Janowski BA, Huffman KE, Schwartz JC, Ram R, Nordsell R, Shames DS, Minna JD, Corey DR. Involvement of AGO1 and AGO2 in mammalian transcriptional silencing. *Nat Struct Mol Biol* 13: 787–792, 2006.
39. Jiang G, Yang F, van Overveld PG, Vedanarayanan V, van der Maarel S, Ehrlich M. Testing the position-effect variegation hypothesis for facioscapulohumeral muscular dystrophy by analysis of histone modification and gene expression in subtelomeric 4q. *Hum Mol Genet* 12: 2909–2921, 2003.
40. Kantor B, Kaufman Y, Makedonski K, Razin A, Shemer R. Establishing the epigenetic status of the Prader-Willi/Angelman imprinting center in the gametes and embryo. *Hum Mol Genet* 13: 2767–2779, 2004.
41. Kaufman Y, Heled M, Perk J, Razin A, Shemer R. Protein-binding elements establish in the oocyte the primary imprint of the Prader-Willi/Angelman syndromes domain. *Proc Natl Acad Sci USA* 106: 10242–10247, 2009.
42. Kim DH, Villeneuve LM, Morris KV, Rossi JJ. Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells. *Nat Struct Mol Biol* 13: 793–797, 2006.
43. Kirschmann DA, Lininger RA, Gardner LM, Seftor EA, Otero VA, Ainsztein AM, Earnshaw WC, Wallrath LL, Hendrix MJ. Down-regulation of HP1Hsalpha expression is associated with the metastatic phenotype in breast cancer. *Cancer Res* 60: 3359–3363, 2000.
44. Kishino T, Lalonde M, Wagstaff J. UBE3A/E6-AP mutations cause Angelman syndrome. *Nat Genet* 15: 70–73, 1997.
45. Kissel JT. Facioscapulohumeral dystrophy. *Semin Neurol* 19: 35–43, 1999.
46. Klooster R, Straasheijm K, Shah B, Sowden J, Frants R, Thornton C, Tawil R, van der Maarel S. Comprehensive expression analysis of FSHD candidate genes at the mRNA and protein level. *Eur J Hum Genet* 17: 1615–1624, 2009.
47. Korber P, Luckenbach T, Blaschke D, Horz W. Evidence for histone eviction in trans upon induction of the yeast PHO5 promoter. *Mol Cell Biol* 24: 10965–10974, 2004.
48. Krol J, Fiszer A, Mykowska A, Sobczak K, de Mezer M, Krzyzosiak WJ. Ribonuclease dicer cleaves triplet repeat hairpins into shorter repeats that silence specific targets. *Mol Cell* 25: 575–586, 2007.
49. Kubicek S, O'Sullivan RJ, August EM, Hickey ER, Zhang Q, Teodoro ML, Rea S, Mechtler K, Kowalski JA, Homon CA, Kelly TA, Jenuwein T. Reversal of H3K9me2 by a small-molecule inhibitor for the G9a histone methyltransferase. *Mol Cell* 25: 473–481, 2007.
50. Kumari D, Usdin K. Chromatin remodeling in the noncoding repeat expansion diseases. *J Biol Chem* 284: 7413–7417, 2009.
51. Kwon MJ, Kim SS, Choi YL, Jung HS, Balch C, Kim SH, Song YS, Marquez VE, Nephew KP, Shin YK. Derepression of CLDN3 and CLDN4 during ovarian tumorigenesis is associated with loss of repressive histone modifications. *Carcinogenesis* 2010 Jan 6. [Epub ahead of print]
52. Lachner M, Sengupta R, Schotta G, Jenuwein T. Trilogies of histone lysine methylation as epigenetic landmarks of the eukaryotic genome. *Cold Spring Harb Symp Quant Biol* 69: 209–218, 2004.
53. Ladd PD, Smith LE, Rabaia NA, Moore JM, Georges SA, Hansen RS, Hagerman RJ, Tassone F, Tapscott SJ, Filippova GN. An antisense transcript spanning the CGG repeat region of FMR1 is upregulated in premutation carriers but silenced in full mutation individuals. *Hum Mol Genet* 16: 3174–3187, 2007.
54. Lakshmikuttyamma A, Scott SA, Decoteau JF, Geyer CR. Reexpression of epigenetically silenced AML tumor suppressor genes by SUV39H1 inhibition. *Oncogene* 29: 576–588, 2010.
55. Lee S, Wevrick R. Identification of novel imprinted transcripts in the Prader-Willi syndrome and Angelman syndrome deletion region: further evidence for regional imprinting control. *Am J Hum Genet* 66: 848–858, 2000.
56. Lehnertz B, Ueda Y, Derijck AA, Braunschweig U, Perez-Burgos L, Kubicek S, Chen T, Li E, Jenuwein T, Peters AH. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr Biol* 13: 1192–1200, 2003.
57. Matsura T, Sutcliffe JS, Fang P, Galjaard RJ, Jiang YH, Benton CS, Rommens JM, Beaudet AL. De novo truncating mutations in E6-AP ubiquitin-protein ligase gene (UBE3A) in Angelman syndrome. *Nat Genet* 15: 74–77, 1997.
58. Meng CF, Zhu XJ, Peng G, Dai DQ. Promoter histone H3 lysine 9 di-methylation is associated with DNA methylation and aberrant expression of p16 in gastric cancer cells. *Oncol Rep* 22: 1221–1227, 2009.
59. Mirkin SM. Expandable DNA repeats and human disease. *Nature* 447: 932–940, 2007.
60. Mito Y, Henikoff JG, Henikoff S. Genome-scale profiling of histone H3.3 replacement patterns. *Nat Genet* 37: 1090–1097, 2005.
61. Montermini L, Andermann E, Labuda M, Richter A, Pandolfo M, Cavalcanti F, Pianese L, Iodice L, Farina G, Monticelli A, Turano M, Filla A, De Michele G, Coccozza S. The Friedreich ataxia GAA triplet repeat: premutation and normal alleles. *Hum Mol Genet* 6: 1261–1266, 1997.
62. Montermini L, Richter A, Morgan K, Justice CM, Julien D, Castellotti B, Mercier J, Poirier J, Capozzoli F, Bouchard JP, Lemieux B, Mathieu J, Vanasse M, Seni MH, Graham G, Andermann F, Andermann E, Melancon SB, Keats BJ, Di Donato S, Pandolfo M. Phenotypic variability in Friedreich ataxia: role of the associated GAA triplet repeat expansion. *Ann Neurol* 41: 675–682, 1997.
63. Montoya-Durango DE, Liu Y, Teneng I, Kalbfleisch T, Lacy ME, Steffen MC, Ramos KS. Epigenetic control of mammalian LINE-1 retrotransposon by retinoblastoma proteins. *Mutat Res* 665: 20–28, 2009.
64. Morris KV, Santoso S, Turner AM, Pastori C, Hawkins PG. Bidirectional transcription directs both transcriptional gene activation and suppression in human cells. *PLoS Genet* 4: e1000258, 2008.
65. Narita M, Nunez S, Heard E, Lin AW, Hearn SA, Spector DL, Hannon GJ, Lowe SW. Rb-mediated heterochromatin formation and silencing of E2F target genes during cellular senescence. *Cell* 113: 703–716, 2003.
66. O'Hagan HM, Mohammad HP, Baylin SB. Double strand breaks can initiate gene silencing and SIRT1-dependent onset of DNA methylation in an exogenous promoter CpG island. *PLoS Genet* 4: e1000155, 2008.
67. Pandolfo M, Pastore A. The pathogenesis of Friedreich ataxia and the structure and function of frataxin. *J Neurol* 256, Suppl 1: 9–17, 2009.
68. Pannetier M, Julien E, Schotta G, Tardat M, Sardet C, Jenuwein T, Feil R. PR-SET7 and SUV4–20H regulate H4 lysine-20 methylation at imprinting control regions in the mouse. *EMBO Rep* 9: 998–1005, 2008.
69. Perk J, Makedonski K, Lande L, Cedar H, Razin A, Shemer R. The imprinting mechanism of the Prader-Willi/Angelman regional control center. *EMBO J* 21: 5807–5814, 2002.
70. Peters AH, O'Carroll D, Scherthan H, Mechtler K, Sauer S, Schofer C, Weipoltshammer K, Pagani M, Lachner M, Kohlmaier A, Opravil S, Doyle M, Sibilia M, Jenuwein T. Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* 107: 323–337, 2001.
71. Phalke S, Nickel O, Walluscheck D, Hortic F, Onorati MC, Reuter G. Retrotransposon silencing and telomere integrity in somatic cells of *Drosophila* depends on the cytosine-5 methyltransferase DNMT2. *Nat Genet* 41: 696–702, 2009.
72. Pirozhkova I, Petrov A, Dmitriev P, Laoudj D, Lipinski M, Vasetzky Y. A functional role for 4qA/B in the structural rearrangement of the 4q35 region and in the regulation of FRG1 and ANT1 in facioscapulohumeral dystrophy. *PLoS One* 3: e3389, 2008.
73. Rai M, Soragni E, Jenssen K, Burnett R, Herman D, Coppola G, Geschwind DH, Gottesfeld JM, Pandolfo M. HDAC inhibitors correct frataxin deficiency in a Friedreich ataxia mouse model. *PLoS One* 3: e1958, 2008.
74. Ricci E, Galluzzi G, Deidda G, Cacurri S, Colantoni L, Merico B, Piazza N, Servidei S, Vigneti E, Pasceri V, Silvestri G, Mirabella M, Mangiola F, Tonali P, Felicetti L. Progress in the molecular diagnosis of facioscapulohumeral muscular dystrophy and correlation between the

- number of KpnI repeats at the 4q35 locus and clinical phenotype. *Ann Neurol* 45: 751–757, 1999.
75. Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, Goodnough LH, Helms JA, Farnham PJ, Segal E, Chang HY. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* 129: 1311–1323, 2007.
 76. Rougeulle C, Cardoso C, Fontes M, Colleaux L, Lalande M. An imprinted antisense RNA overlaps UBE3A and a second maternally expressed transcript. *Nat Genet* 19: 15–16, 1998.
 77. Saitoh S, Wada T. Parent-of-origin specific histone acetylation and reactivation of a key imprinted gene locus in Prader-Willi syndrome. *Am J Hum Genet* 66: 1958–1962, 2000.
 78. Sasai N, Defossez PA. Many paths to one goal? The proteins that recognize methylated DNA in eukaryotes. *Int J Dev Biol* 53: 323–334, 2009.
 79. Saveliev A, Everett C, Sharpe T, Webster Z, Festenstein R. DNA triplet repeats mediate heterochromatin-protein-1-sensitive variegated gene silencing. *Nature* 422: 909–913, 2003.
 80. Schotta G, Ebert A, Dorn R, Reuter G. Position-effect variegation and the genetic dissection of chromatin regulation in *Drosophila*. *Semin Cell Dev Biol* 14: 67–75, 2003.
 81. Schotta G, Lachner M, Sarma K, Ebert A, Sengupta R, Reuter G, Reinberg D, Jenuwein T. A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev* 18: 1251–1262, 2004.
 82. Schotta G, Sengupta R, Kubicek S, Malin S, Kauer M, Callen E, Celeste A, Pagani M, Opravil S, De La Rosa-Velazquez IA, Espejo A, Bedford MT, Nussenzweig A, Busslinger M, Jenuwein T. A chromatin-wide transition to H4K20 monomethylation impairs genome integrity and programmed DNA rearrangements in the mouse. *Genes Dev* 22: 2048–2061, 2008.
 83. Sharma S, Kelly TK, Jones PA. Epigenetics in cancer. *Carcinogenesis* 31: 27–36, 2010.
 84. Tan J, Yang X, Zhuang L, Jiang X, Chen W, Lee PL, Karuturi RK, Tan PB, Liu ET, Yu Q. Pharmacologic disruption of Polycomb-repressive complex 2-mediated gene repression selectively induces apoptosis in cancer cells. *Genes Dev* 21: 1050–1063, 2007.
 85. Tawil R, Forrester J, Griggs RC, Mendell J, Kissel J, McDermott M, King W, Weiffenbach B, Figlewicz D. Evidence for anticipation and association of deletion size with severity in facioscapulohumeral muscular dystrophy. The FSH-DY Group. *Ann Neurol* 39: 744–748, 1996.
 86. Trojer P, Reinberg D. Facultative heterochromatin: is there a distinctive molecular signature? *Mol Cell* 28: 1–13, 2007.
 87. Tryndyak VP, Kovalchuk O, Pogribny IP. Loss of DNA methylation and histone H4 lysine 20 trimethylation in human breast cancer cells is associated with aberrant expression of DNA methyltransferase 1, Suv4–20h2 histone methyltransferase and methyl-binding proteins. *Cancer Biol Ther* 5: 65–70, 2006.
 88. Van Den Broeck A, Brambilla E, Moro-Sibilot D, Lantuejoul S, Brambilla C, Eymin B, Khochbin S, Gazzeri S. Loss of histone H4K20 trimethylation occurs in preneoplasia and influences prognosis of non-small cell lung cancer. *Clin Cancer Res* 14: 7237–7245, 2008.
 89. van Deutekom JC, Wijmenga C, van Tienhoven EA, Gruter AM, Hewitt JE, Padberg GW, van Ommen GJ, Hofker MH, Frants RR. FSHD associated DNA rearrangements are due to deletions of integral copies of a 3.2 kb tandemly repeated unit. *Hum Mol Genet* 2: 2037–2042, 1993.
 90. van Koningsbruggen S, Dirks RW, Mommaas AM, Onderwater JJ, Deidda G, Padberg GW, Frants RR, van der Maarel SM. FRG1P is localised in the nucleolus, Cajal bodies, and speckles. *J Med Genet* 41: e46, 2004.
 91. van Koningsbruggen S, Straasheijm KR, Sterrenburg E, de Graaf N, Dauwerse HG, Frants RR, van der Maarel SM. FRG1P-mediated aggregation of proteins involved in pre-mRNA processing. *Chromosoma* 116: 53–64, 2007.
 92. van Overveld PG, Lemmers RJ, Sandkuijl LA, Enthoven L, Winokur ST, Bakels F, Padberg GW, van Ommen GJ, Frants RR, van der Maarel SM. Hypomethylation of D4Z4 in 4q-linked and non-4q-linked facioscapulohumeral muscular dystrophy. *Nat Genet* 35: 315–317, 2003.
 93. Varga-Weisz PD, Becker PB. Regulation of higher-order chromatin structures by nucleosome-remodelling factors. *Curr Opin Genet Dev* 16: 151–156, 2006.
 94. Wells RD. DNA triplexes and Friedreich ataxia. *FASEB J* 22: 1625–1634, 2008.
 95. Williams CA, Beaudet AL, Clayton-Smith J, Knoll JH, Kyllerman M, Laan LA, Magenis RE, Moncla A, Schinzel AA, Summers JA, Wagstaff J. Angelman syndrome 2005: updated consensus for diagnostic criteria. *Am J Med Genet A* 140: 413–418, 2006.
 96. Wu MY, Tsai TF, Beaudet AL. Deficiency of Rbbp1/Arid4a and Rbbp111/Arid4b alters epigenetic modifications and suppresses an imprinting defect in the PWS/AS domain. *Genes Dev* 20: 2859–2870, 2006.
 97. Wu ZY, Wang ZQ, Murong SX, Wang N. FSHD in Chinese population: characteristics of translocation and genotype-phenotype correlation. *Neurology* 63: 581–583, 2004.
 98. Xin Z, Allis CD, Wagstaff J. Parent-specific complementary patterns of histone H3 lysine 9 and H3 lysine 4 methylation at the Prader-Willi syndrome imprinting center. *Am J Hum Genet* 69: 1389–1394, 2001.
 99. Xin Z, Tachibana M, Guggiari M, Heard E, Shinkai Y, Wagstaff J. Role of histone methyltransferase G9a in CpG methylation of the Prader-Willi syndrome imprinting center. *J Biol Chem* 278: 14996–15000, 2003.
 100. Zeng W, de Greef JC, Chen YY, Chien R, Kong X, Gregson HC, Winokur ST, Pyle A, Robertson KD, Schmiesing JA, Kimonis VE, Balog J, Frants RR, Ball AR Jr, Lock LF, Donovan PJ, van der Maarel SM, Yokomori K. Specific loss of histone H3 lysine 9 trimethylation and HP1gamma/cohesin binding at D4Z4 repeats is associated with facioscapulohumeral dystrophy (FSHD). *PLoS Genet* 5: e1000559, 2009.
 101. Zhao J, Sun BK, Erwin JA, Song JJ, Lee JT. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* 322: 750–756, 2008.

2.2 Epigenetic regulation of development by histone lysine methylation.

Review article: Dambacher et al. 2010, Heredity

REVIEW

Epigenetic regulation of development by histone lysine methylation

S Dambacher¹, M Hahn¹ and G SchottaMunich Center for Integrated Protein Science (CiPS^M) and Adolf-Butenandt-Institute, Ludwig-Maximilians-University, Munich, Germany

Epigenetic mechanisms contribute to the establishment and maintenance of cell-type-specific gene expression patterns. In this review, we focus on the functions of histone lysine methylation in the context of epigenetic gene regulation during developmental transitions. Over the past few years, analysis of histone lysine methylation in active and repressive nuclear compartments and, more recently, genome-wide profiling of histone lysine methylation in different cell types have revealed correlations between particular modifications and the transcriptional status of genes. Identification of histone methyltransferases

(HMTases) and specific binding factors for most methylated lysine positions has provided a novel insight into the mechanisms of epigenetic gene regulation. In addition, analyses of HMTase knockout mice show that histone lysine methylation has important functions for normal development. In this study, we review mechanisms of gene activation and repression by histone lysine methylation and discuss them in the context of the developmental roles of HMTases.

Heredity (2010) **105**, 24–37; doi:10.1038/hdy.2010.49; published online 5 May 2010

Keywords: epigenetics; histone lysine methylation; heterochromatin; mouse development

Introduction

Development is accomplished by spatial and temporal regulation of gene expression patterns. The identity of each cell type is maintained and passed on to daughter cells by mechanisms that do not alter the DNA sequence and are therefore regarded as epigenetic. A major mechanism to establish cell-type-specific expression patterns is transcriptional regulation. The physiological template for transcription is chromatin, and therefore epigenetic mechanisms are thought to modulate its structure, making DNA more or less accessible to the transcriptional machinery. Today, we know of five major mechanisms that alter chromatin architecture: DNA methylation, post-translational histone modifications, use of histone variants, chromatin remodeling and incorporation of non-coding RNA into chromatin. These mechanisms are generally considered ‘epigenetic’, although we still lack good understanding as to the stability of these modifications through mitosis or even through the germ line.

In this review, we will focus on the functions of histone lysine methylation during development. In the first part, we will discuss how histone lysine methylation facilitates gene activation or repression of genomic regions. In the context of these activities, we will then discuss the developmental roles of selected histone methyltransferases (HMTases).

Activation and repression are facilitated by histone lysine methylation

Major methylation sites on histones H3 and H4 are located in the tail (H3K4, H3K9, H3K27, H3K36 and H4K20) and the nucleosome core region (H3K79). Although histone methylation was discovered nearly four decades ago (Allfrey *et al.*, 1964), a correlation between this modification and gene regulation has only recently been established. Strahl *et al.* (1999) showed that H3K4 methylation was highly enriched in macronuclei of *Tetrahymena*, suggesting a role for this modification in transcriptional activation. A year later, H3K9 methylation was implicated in gene repression when a homolog of the heterochromatin-associated *Drosophila* Su(var)3-9, Suv39h1, was shown to have H3K9-specific methyltransferase activity (Rea *et al.*, 2000).

Lysine residues can be mono (me1), di (me2) or trimethylated (me3), and binding of specific proteins, which recognize methylated lysine positions, can result in different biological outcomes. The development of highly specific antibodies that discriminate not only between lysine positions, but also between methylation states, allowed the large-scale mapping of individual histone lysine methylation marks by chromatin immunoprecipitation on tiling arrays or chromatin immunoprecipitation followed by sequencing (reviewed in Lee and Mahadevan, 2009). In combination with gene expression data, correlations between histone modifications and gene activity have now been established. A number of modifications show some correlation with the transcriptional status of genes; however, only a very few marks are consistently found on active or inactive genes. The hallmark of transcriptionally active genes is H3K4me3 in the promoter region and H3K36me3

Correspondence: Professor G Schotta, Adolf-Butenandt-Institute, Ludwig-Maximilians-University, Schillerstrasse 44, Munich 80336, Germany.

E-mail: gunnar.schotta@med.uni-muenchen.de

¹These authors contributed equally to this work.

Received 8 November 2009; revised 16 March 2010; accepted 18 March 2010; published online 5 May 2010

across the gene body (Figure 1). H3K27me3 seems central for gene repression and covers the gene body and flanking regions (Figure 1). The average profile of repressed genes shows enrichment of H3K9me3 and H4K20me3 at a much lower level when compared with H3K27me3, suggesting that these marks are less important for gene silencing (Figure 1). In the next sections, we will discuss the current view of how active and repressive modifications are established and how they contribute to the transcriptional regulation of genes.

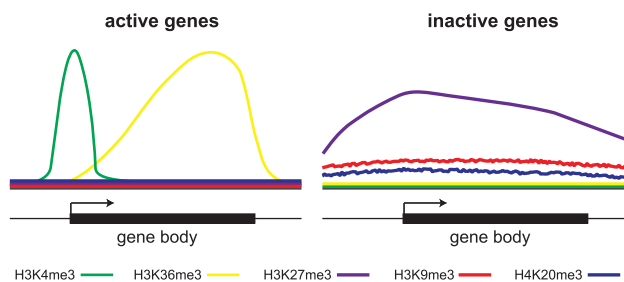


Figure 1 Profiles of histone lysine methylation at active and inactive mammalian genes. At active genes, H3K4me3 is highly enriched at the promoter region, whereas H3K36me3 associates with the gene body. At repressed genes, H3K27me3 covers the gene body and flanking regions. H3K9me3 and H4K20me3 are also enriched at inactive genes, although at a lower level (adapted from Barski *et al.*, 2007, Pauler *et al.*, 2009 and Cui *et al.*, 2009).

Histone lysine methylation marks in the context of transcriptional activation

The amount of transcript per cell is controlled through multiple mechanisms. If we only consider the rate of transcription, at least two major steps regulate how much primary transcript is produced. In the first step, RNA polymerase II (Pol II) is recruited to the promoter region and forms a pre-initiation complex (PIC). However, this is not sufficient for transcription, as RNA Pol II can be stalled at promoters and a second trigger is therefore required for elongation (Core and Lis, 2008). Histone lysine methylation is an important regulatory element for both determining processes. It exerts an effect by recruiting specific binding factors, providing stable interaction platforms for the basic transcriptional machinery and for activities that regulate the ordered dis- and reassembly of chromatin during elongation.

Transcription initiation

Active genes carry high levels of H3K4me3 in the promoter region. This modification is a binding platform for a number of proteins (Table 1), including chromatin remodelers, which help to open the chromatin structure around the promoter and facilitate the binding of the basic transcriptional machinery. BPTF (bromodomain PHD finger transcription factor), a subunit of the NURF (nucleosome remodeling factor) remodeling complex, and Chd1, another chromatin

Table 1 Mammalian histone methylation binders and their possible functions

Histone modification	Reader	Function	Reference
H3K4			
me2, me3	Chd1	Chromatin remodeling	Flanagan <i>et al.</i> (2005)
me3	Bptf	Chromatin remodeling	Wysocka <i>et al.</i> (2006)
me3	Taf3	TFIID stabilization	Vermeulen <i>et al.</i> (2007)
me3	Ing1	Recruitment of HATs	Taverna <i>et al.</i> (2006)
me3	Ing2	Recruitment of HDACs	Shi <i>et al.</i> (2006)
me3	Ing4	Recruitment of HATs	Hung <i>et al.</i> (2009)
me3	Ing5	Recruitment of HATs	Champagne <i>et al.</i> (2008)
me3	Jmjd2a	H3K9 demethylation	Lee <i>et al.</i> (2008b)
me3	Chd7	Chromatin remodeling	Takada <i>et al.</i> (2007)
me3	Rag2	VDJ recombination	Matthews <i>et al.</i> (2007)
H3K9			
me1, me2	G9a-GLP	Transcriptional silencing	Collins <i>et al.</i> (2008)
me3	HP1	Heterochromatin	Lachner <i>et al.</i> (2001)
me3	Tip60	DNA repair	Sun <i>et al.</i> (2009)
me3	Chd7	Chromatin remodeling	Takada <i>et al.</i> (2007)
me3	Cdyl2	Heterochromatin	Fischle <i>et al.</i> (2008)
H3K27			
me3	Cbx2,4,7	Polycomb-mediated gene silencing	Bernstein <i>et al.</i> (2006b)
me3	Eed	Polycomb-mediated gene silencing	Margueron <i>et al.</i> (2009)
H3K36			
me3	Mrg15	Recruitment of HDACs	Zhang <i>et al.</i> (2006)
H3K79			
me1, me2	53bp1	DNA damage repair	Huyen <i>et al.</i> (2004)
H4K20			
me1, me2	L3mbtl1	Chromatin compaction	Trojer <i>et al.</i> (2007)
me2	53bp1	DNA damage repair	Botuyan <i>et al.</i> (2006)
me3	Jmjd2a	H3K9 demethylation	Lee <i>et al.</i> (2008b)

Abbreviations: HAT, histone acetyltransferase; HDAC, histone deacetylase.

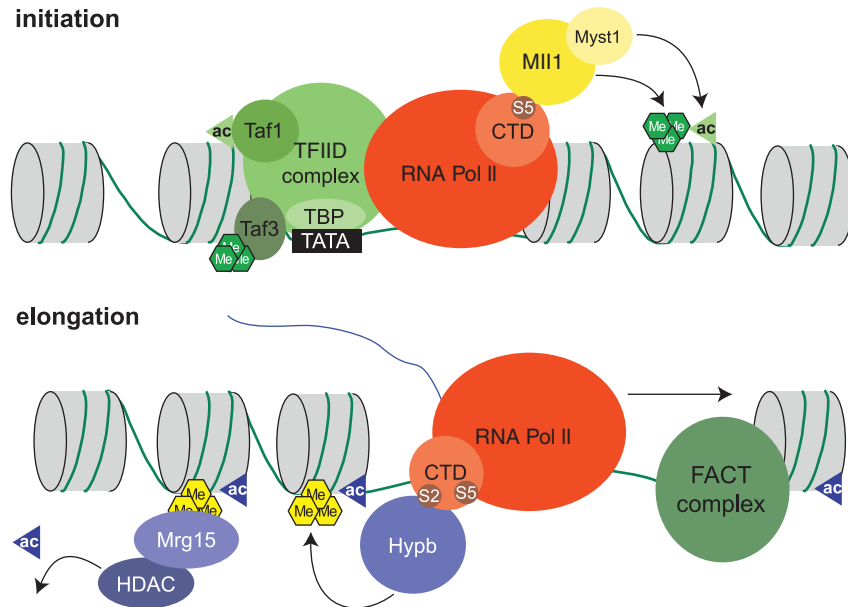


Figure 2 Histone lysine methylation marks in the context of transcriptional activation. During initiation, TFIID is targeted to the promoter region through multiple interactions between its subunits and chromatin modifications (see text for details). RNA Pol II is positioned near the transcriptional start site and phosphorylated at Ser 5 of its CTD. The Mll1 complex binds to S5-P CTD and introduces H3K4me3 as well as acetylation of H4. Phosphorylation at Ser 2 of the CTD starts the elongation phase. The FACT complex disrupts nucleosomes in front of elongating RNA Pol II. Hypb binds to S2-P CTD and induces H3K36me3, which is recognized by Mrg15 leading to deacetylation of histones due to recruitment of histone deacetylases (HDACs).

remodeler, can bind H3K4me3 (Flanagan *et al.*, 2005; Wysocka *et al.*, 2006). Although these data implicate chromatin remodeling in promoter activation, we still lack detailed mechanistic insight.

Other binding proteins of H3K4me3 seem to be important for recognition of the promoter region. The TFIID complex is involved in the first step of PIC formation, and binds the promoter through multiple interactions between its subunits, DNA and histone modifications (Figure 2). The TATA box binding protein (TBP) subunit, for example, and other associated proteins recognize the promoter DNA sequence. This can be either the TATA box itself or associated sequences, such as initiator and downstream promoter elements. The chromatin state is recognized by another TFIID subunit, the double bromodomain protein Taf1, which can bind acetylated lysines at positions H3K9 and H3K14 (Jacobson *et al.*, 2000). More recently, yet another TFIID component, the PHD domain protein Taf3, was shown to bind H3K4me3 (Vermeulen *et al.*, 2007). These data suggest that multiple interactions are necessary for stable recruitment of the PIC. The next step in the initiation cascade is binding of RNA Pol II to the promoter region and phosphorylation of Ser 5 in its carboxy-terminal domain (CTD) repeats by Cdk7. RNA Pol II is then able to generate short transcripts; however, interaction with DSIF (DRB sensitivity-inducing factor) and NELF (negative elongation factor) can inhibit elongation (reviewed in Fuda *et al.*, 2009).

Is H3K4me3 at the promoter a cause or a consequence of transcriptional activity? Currently, there is no general answer to this question and every promoter might behave somewhat differently. *In vitro* data suggest that this modification has no direct effect on transcription (Pavri *et al.*, 2006; Kim and Buratowski, 2009); however, the reduced complexity of *in vitro* systems might mask

an effect that is relevant *in vivo*. It is noteworthy that some inducible promoters carry H3K4me3 even before RNA Pol II is detectable, indicating that H3K4me3 can be established in the absence of active transcription (Edmunds *et al.*, 2008). Several enzymes can induce H3K4me3 (Table 2), and elucidating how they are recruited and how their HMTase activity is regulated will be important in furthering our understanding of this modification. There is evidence that the CTD exerts an effect as a recruiting platform for different chromatin-modifying activities. For example, Mll1, a major H3K4-specific methyltransferase, interacts with the Ser 5 phosphorylated CTD and establishes or reinforces H3K4me3 around the promoter region (Milne *et al.*, 2005).

Transcription elongation

Binding to H3K4me3 might stabilize the PIC at the promoter and could therefore increase the probability of initiating transcription. The other mechanism to control the transcription rate is elongation (Figure 2). RNA Pol II can be stalled at promoters by interaction with NELF and DSIF. Elongation is then induced by phosphorylation of DSIF and RNA Pol II CTD at Ser 2 by the P-TEFb (positive transcription elongation factor b) complex (reviewed in Fuda *et al.*, 2009). For RNA Pol II to traverse nucleosomal templates, the chromatin structure needs to be relaxed. This is facilitated through eviction of H2A/H2B dimers by the FACT complex (Belotserkovskaya and Reinberg, 2004). After passage of RNA Pol II, the FACT complex could also be involved in the reassembly of a proper chromatin structure (Jamai *et al.*, 2009).

Several histone modifications are established when RNA Pol II travels through the gene body. The Ser 2 phosphorylated CTD associates with H3K36-specific

Table 2 Mammalian HMTases, their activities and knockout phenotypes

HMTase	Activity	Reference	Viability	Phenotype
Ash1l Dot1l	H3K4 H3K79	Gregory <i>et al.</i> (2007) Feng <i>et al.</i> (2002)	E9.5–10.5	Growth retardation, angiogenesis defects in the yolk sac, and cardiac dilation, loss of all H3K79 methylation (Jones <i>et al.</i> , 2008)
Ezh1	H3K27	Margueron <i>et al.</i> (2008); Shen <i>et al.</i> (2008)		
Ezh2	H3K27	Cao <i>et al.</i> (2002)	E8.5	Arrested development, gastrulation failure (O'Carroll <i>et al.</i> 2001)
G9a	H3K9me1/2	Tachibana <i>et al.</i> (2002)	E9.5–12.5	Growth retardation, reduction in H3K9 me1, me2 (Tachibana <i>et al.</i> , 2002)
GLP	H3K9me1/2	Tachibana <i>et al.</i> (2008)	E9.5–12.5	Growth retardation, reduction in H3K9 me1, me2 (Tachibana <i>et al.</i> , 2005)
MLL1	H3K4	Milne <i>et al.</i> (2002)	E12.5–16.5	Patterning defects (Yu <i>et al.</i> , 1995; Yagi <i>et al.</i> , 1998) Δ SET mutant viable, skeletal defects (Terranova <i>et al.</i> , 2006)
MLL2	H3K4	Goo <i>et al.</i> (2003)	E11.5	Growth retardation, increased apoptosis (Glaser <i>et al.</i> , 2006)
MLL3	H3K4	Lee <i>et al.</i> (2006)	Viable	Partial embryonic lethality, growth retardation, female infertility (Lee <i>et al.</i> , 2006)
MLL4	H3K4	Lee <i>et al.</i> (2006)		
MLL5	H3K4	Fujiki <i>et al.</i> (2009)	Viable	Hematopoietic defects, male infertility (Heuser <i>et al.</i> , 2009; Liu <i>et al.</i> , 2009; Madan <i>et al.</i> , 2009)
Nsd1	H3K36	Rayasam <i>et al.</i> (2003)	E9.5	Growth retardation, apoptosis (Rayasam <i>et al.</i> , 2003)
Prdm1			E10.5	Patterning defects (Ohinata <i>et al.</i> , 2005)
Prdm2	H3K9me2	Kim <i>et al.</i> (2003)	Viable	Tumorigenesis (Steele-Perkins <i>et al.</i> , 2001)
Prdm3			E13.5–16.5	Broad developmental defects (Hoyt <i>et al.</i> , 1997)
Prdm4				
Prdm5				
Prdm6				
Prdm8				
Prdm9	H3K4me3	Hayashi <i>et al.</i> (2005)	Viable	Impaired sex body formation, infertility (Hayashi <i>et al.</i> , 2005)
Prdm10				
Prdm11				
Prdm12				
Prdm13				
Prdm14			Viable	Infertility, devoid of germ cells (Yamaji <i>et al.</i> , 2008)
Prdm15				
Prdm16				
PrSet7	H4K20me1	Nishioka <i>et al.</i> (2002)	Eight-cell stage	G2/M arrest, chromosome condensation defects (Oda <i>et al.</i> , 2009)
Setd1a	H3K4	Wysocka <i>et al.</i> (2003)		
Setd1b	H3K4	Lee <i>et al.</i> (2007)		
Setd2	H3K36	Sun <i>et al.</i> (2005)		
Setd3				
Setd4				
Setd5				
Setd6				
Setd7	H3K4me1/2	Wang <i>et al.</i> (2001)	Viable	50% embryonic lethality (Kurash <i>et al.</i> , 2008)
Setdb1	H3K9	Yang <i>et al.</i> (2002)	E3.5–5.5	Defective growth of inner cell mass (Dodge <i>et al.</i> , 2004)
Setdb2				
Setmar	H3K36	Lee <i>et al.</i> (2005)		
Smyd1	H3K4me1/2/3	Tan <i>et al.</i> (2006)	E10.5	Growth retardation, disrupted maturation of ventricular cardiomyocytes (Gottlieb <i>et al.</i> , 2002)
Smyd2	H3K36	Brown <i>et al.</i> (2006)		
Smyd3	H3K4me2/3	Hamamoto <i>et al.</i> (2004)		
Smyd4				
Smyd5				
Suv39h1	H3K9me2/3	Rea <i>et al.</i> (2000)	Viable	Suv39h dn mice: growth retardation, increased tumor risk (B-cell lymphomas), male sterility (Peters <i>et al.</i> , 2001)
Suv39h2	H3K9me2/3	Rea <i>et al.</i> (2000)	Viable	
Suv4-20h1	H4K20me2/3	Schotta <i>et al.</i> (2004)	Perinatal lethality	Suv4-20h dn mice: growth retardation, lung defects, impaired B-cell development (Schotta <i>et al.</i> , 2008)
Suv4-20h2	H4K20me2/3	Schotta <i>et al.</i> (2004)	Viable	
Whsc1	H3K36	Nimura <i>et al.</i> (2009)	Perinatal lethality	Growth retardation, defects in midline fusion, heart lesions (Nimura <i>et al.</i> , 2009)
Whsc1l1				

Abbreviations: dn, double null; HMTase, histone methyltransferase.

HMTases during elongation (Yoh *et al.*, 2008). H3K36me3 generates a binding platform for the chromo domain protein Mrg15 (Table 1, Zhang *et al.*, 2006), which in turn recruits histone deacetylases (Yochum and Ayer, 2002). H3K36me3 and histone deacetylation are important to repress transcripts that could be generated from aberrant

initiation of RNA Pol II within the gene body (Carrozza *et al.*, 2005). Thus, histone modifications in the context of elongation might indirectly affect the elongation rate and have important functions for re-establishment of a proper chromatin structure during and after transcription. Interestingly, recent data suggest that histone

modifications across the gene body can even affect processing of the RNA transcript, such as selection of polyadenylation sites or even splicing (Kolasinska-Zwiercz *et al.*, 2009; Spies *et al.*, 2009). These surprising connections between chromatin structure and RNA processing will reveal novel mechanisms for the regulation of cell-type-specific transcription profiles.

Repressive histone lysine methylation marks

Transcriptional repression is important for various aspects of development. On one hand, cell-type-specific expression patterns are regulated by silencing of lineage-inappropriate genes during differentiation, and on the other hand, large regions of mammalian genomes consist of non-coding DNA sequences such as satellite repeats, telomeric repeats, mobile elements and interspersed repeats, which need to be under tight transcriptional control to prevent genomic instability. Genome-wide mapping studies of histone modifications consider only the nonrepetitive part of the genome and from these data it seems that H3K27me3 is a major modification that correlates with the transcriptional repression of genes (Figure 1). In contrast, repetitive genomic regions are marked by H3K9me3 and H4K20me3. In the following sections, we will discuss the establishment and potential functions of repressive histone modifications at both highly repetitive genomic regions and at individual genes.

Silencing of repetitive genomic regions

Pericentric heterochromatin

The largest family of repetitive regions consists of major satellite repeats that are the main constituents of pericentric heterochromatin. Major satellite repeats have a distinct H3K9me3 + H4K20me3 chromatin signature, which is found in almost all cell types and developmental stages, suggesting that these modifications have a general function in heterochromatin. H3K9me3 is established by Suv39h1 and Suv39h2 enzymes (Rea *et al.*, 2000). Two other HMTases, Suv4-20h1 and Suv4-20h2, establish H4K20me3 (Schotta *et al.*, 2004). Recent data suggest a novel modification, H3K64me3 that has a role in pericentric heterochromatin formation during the early stages of mouse development (Daujat *et al.*, 2009); however, the responsible methyltransferase has not been identified as yet.

The combinatorial pattern of histone lysine methylation at heterochromatin is established in a sequential pathway (Figure 3). In *Schizosaccharomyces pombe*, double-stranded RNA from centromeric repeats is processed by components of the RNA interference machinery. This leads to recruitment of Clr4, the *S. pombe* homolog of Suv39h, to establish H3K9 methylation at heterochromatin (Grewal and Jia, 2007). In mammals, it is still unclear whether a similar link between processing of double-stranded RNA and recruitment of Suv39h exists. H3K9me3 is likely to be established in a highly coordinated manner during replication of pericentric heterochromatin. Recent data suggest that, in a first step, Setdb1 (SET domain bifurcated 1), in complex with heterochromatin protein 1 α (HP1 α) and CAF1 (chromatin assembly factor 1), induces H3K9me1 on

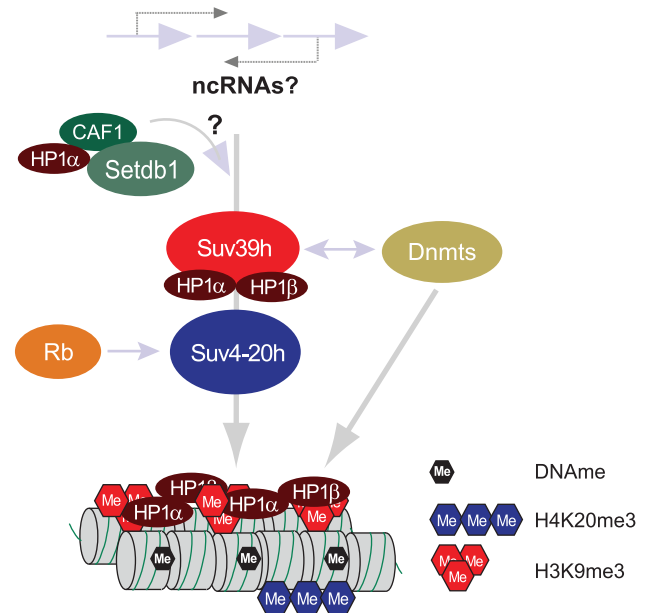


Figure 3 Repressive histone lysine methylation marks at heterochromatin are established in a sequential pathway. The Setdb1/CAF1/HP1 α complex presumably induces H3K9me1, which is converted to H3K9me3 by Suv39h enzymes. H3K9me3 exerts an effect as a binding platform for HP1 proteins, which in turn recruit Suv4-20h enzymes to induce H4K20me3. Establishment of these modifications is also regulated through interactions with other proteins, for example, members of the retinoblastoma (Rb) family. There is also evidence for interaction of DNA methyltransferases with Suv39h enzymes; however, this interplay needs further characterization.

non-nucleosomal histone H3 (Loyola *et al.*, 2009). Subsequently, Suv39h enzymes, which prefer H3K9me1 as substrate, induce H3K9me3, probably even before the H3 molecule is deposited into a nucleosomal context (Rea *et al.*, 2000). Nucleosomal H3K9me3 exerts an effect as a binding platform for HP1 proteins, which in turn recruit Suv4-20h enzymes to establish H4K20me3 (Schotta *et al.*, 2004). The direct interaction of Suv4-20h enzymes with HP1 is necessary to induce H4K20me3; however, interactions with other proteins, for example, members of the retinoblastoma family, can contribute to the establishment of this modification (Gonzalo *et al.*, 2005).

What are the functions of histone lysine methylation marks at pericentric heterochromatin? Although considered highly compact and transcriptionally silent, there is increasing evidence for controlled transcription across pericentric heterochromatin (Eymery *et al.*, 2009). Promoter elements are still unknown, and owing to the repetitive nature of these transcripts they have yet to be characterized. However, it is clear that transcription from major satellite repeats is tightly controlled, occurring only during distinct cell cycle stages. Suv39h double-null mutants show enhanced amounts of major satellite transcripts (Martens *et al.*, 2005), indicating that H3K9me3 has an important role in controlling the transcript levels from these repeat regions. How this control is accomplished, whether H3K9me3 or its binding factors hinder access to RNA Pol II or whether RNA processing or RNA stability are regulated by this modification are some challenging questions in this field.

Telomeric silencing

The chromatin structure at telomeres is very similar to that of pericentric heterochromatin. Telomeric repeats are enriched for Suv39h-mediated H3K9me3. As in heterochromatin, H3K9me3 exerts an effect as a binding platform for HP1 proteins that recruit Suv4-20h enzymes to induce H4K20me3 (Benetti *et al.*, 2007). At telomeres, this sequential pathway is affected by the H3K79-specific HMTase Dot1l. It is noteworthy that Dot1l-mutant cells lose all H3K79 methylation and that at telomeres even H4K20me3 is lost (Jones *et al.*, 2008). Currently, it is not known whether Dot1l or H3K79 methylation affects the activity or recruitment of Suv4-20h enzymes.

Transcripts from telomeric repeats (TelRNAs) are generated by RNA Pol II and are normally polyadenylated (Schoeftner and Blasco, 2008). Interestingly, TelRNAs seem to be a structural constituent of telomeric chromatin. They can block telomerase activity, and therefore a possible function of these transcripts is the regulation of telomere length. In the absence of Suv39h or Suv4-20h enzymes, TelRNAs are upregulated, suggesting that H3K9me3 and H4K20me3 function as repressive modifications at telomeres (Schoeftner and Blasco, 2008).

Gene silencing

Gene activation is largely correlated with the establishment of H3K4me3. In contrast, different modifications exist to mediate gene silencing. Average chromatin immunoprecipitation profiles across silenced genes suggest H3K27me3 as a prominent modification for gene repression (Figure 1); however, there is also evidence that H3K9me3 and H4K20me3 are associated with repressed genes. Interestingly, different studies have found that there is a large proportion of silent genes that do not carry any of the tested epigenetic modifications. It is possible that these genes are passively repressed and that their silent state is just due to the lack of activating factors. However, we still lack knowledge about many histone modifications and their mechanisms of action, and it could well be that novel mechanisms for transcriptional silencing will be discovered in the near future. In the next sections, we will discuss how gene repression is established by the classic repressive histone modifications, H3K27me3 and H3K9me3.

Polycomb silencing by H3K27 methylation

H3K27 methylation only exists in multicellular organisms and has probably evolved as a system to facilitate cell-type differentiation. Surprisingly, in embryonic stem (ES) cells, H3K27me3 can coexist in the same region with H3K4me3 (Bernstein *et al.*, 2006a). Genes that carry this 'bivalent' modification are mainly developmental regulators. Although bivalent genes are repressed, they carry engaged but stalled RNA Pol II. In differentiated cells, bivalent chromatin domains are reduced and genes that are active or repressed are characterized by H3K4me3 or H3K27me3, respectively (Mikkelsen *et al.*, 2007). It is not clear how bivalent marks are reduced to a univalent form during differentiation. One possibility is that during replication the bivalent modification cannot be 'copied' and either H3K4me3 or H3K27me3 is established on the newly replicated chromatin. Another

postulation is that chromatin modifications at a particular gene reflect an equilibrium between antagonizing activities of transcriptional activators and repressors. Unequal cell division during differentiation of pluripotent progenitor cells might shift this balance to either side. Bivalent chromatin might also represent a transient state during differentiation. In ES cells, the major pluripotency genes Nanog, Sox2 and Oct4, are highly expressed and marked with H3K4me3. During differentiation, these genes transiently acquire a bivalent state before they become silenced with H3K27me3 (Pan *et al.*, 2007).

H3K27 methylation is mediated by the two highly related enzymes, Ezh1 (enhancer of zeste homolog 1) and Ezh2 (Table 2). Ezh enzymes form the so-called polycomb repressive complex 2 (PRC2) with Eed, Suz12 and RbAp46/48 proteins (Margueron *et al.*, 2008). It is noteworthy that ES cells in which PRC2 complex members such as Eed or Suz12, are disrupted, largely lose H3K27me2 and H3K27me3 (Schoeftner *et al.*, 2006; Chamberlain *et al.*, 2008), suggesting that these proteins exert an effect as cofactors to fully stimulate enzymatic activity of Ezh enzymes.

H3K27me3 recruits the PRC1 complex (Ring1a/b, Bmi1, Ph, Cbx2) through interaction with the Cbx2 chromo domain (Figure 4). Other Cbx2 homologs bind H3K27me3 and can be part of PRC1-related complexes (Bernstein *et al.*, 2006b). For the recruitment of PRC1, binding to H3K27me3 is essential but probably not sufficient. This became clear through genome-wide mapping of PRC1 and PRC2 complexes (Ku *et al.*, 2008). Only a subset of regions with high levels of H3K27me3 also shows enrichment for PRC1 complex members. However, removal of H3K27me3 leads to complete loss of PRC1 from its targets (Cao *et al.*, 2002; Leeb *et al.*, 2010). An important function of PRC1 is the establishment of a second histone modification, H2A ubiquitylation on

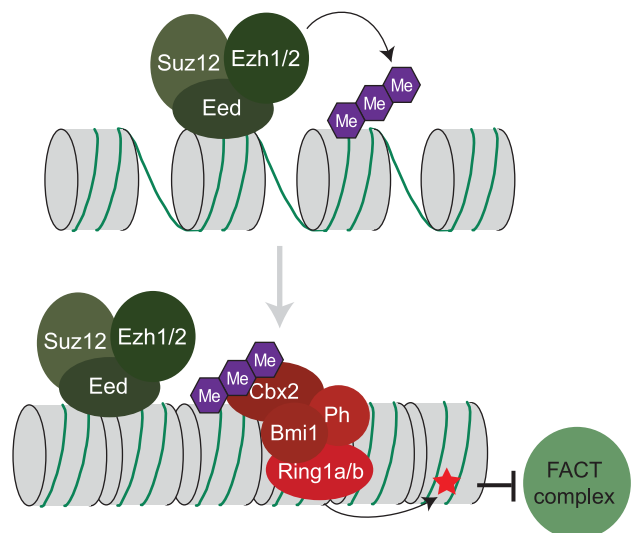


Figure 4 Polycomb-mediated gene silencing. Ezh1/2 enzymes, Suz12 and Eed, form the PRC2 core complex and induce H3K27me3. This modification is recognized by the chromo domain of mammalian Pc homologs, for example Cbx2, which is a subunit of the PRC1 complex. Ring1a/b, another PRC1 subunit, establishes H2AK119ub, which inhibits nucleosome disassembly by the FACT complex.

lysine 119 (H2AK119ub) by the Ring1a/b subunit (Wang *et al.*, 2004).

The mechanism for transcriptional repression by the polycomb system (Figure 4) is still under debate. *In vitro* data suggest that PRC1 and PRC2 components can compact recombinant nucleosomes and block transcriptional elongation on chromatinized templates (Margueron *et al.*, 2008); however, these mechanisms are very difficult to verify *in vivo*. A more compact chromatin structure might as well prevent promoter recognition by the PIC. Another intriguing finding is that H2AK119ub prevents recruitment of FACT (Zhou *et al.*, 2008), which could impair transcriptional elongation. These and maybe other epigenetic mechanisms, such as binding of non-coding RNAs (Rinn *et al.*, 2007) and DNA methylation (Vire *et al.*, 2006), work together in the establishment of polycomb-mediated gene silencing. However, future studies are needed to dissect the interplay between these mechanisms.

Imprinting

Genomic imprinting is a well-characterized system that mainly uses epigenetic mechanisms to induce stable gene repression. Imprinted genes are only expressed from one allele; the other allele is permanently silenced using a heterochromatin-like mechanism. Silencing is established by repressive histone modifications and DNA methylation over so-called imprinting control regions (ICRs), which inhibit promoter activation or can block enhancer action.

Imprinted loci show allelic differences in epigenetic patterns. The ICR of the silenced allele carries heterochromatin-like modifications (H3K9me3 + H4K20me3), whereas the active allele is marked with H3K4me3 (Fournier *et al.*, 2002; Regha *et al.*, 2007; Pannetier *et al.*, 2008). Not all imprinted genes are regulated in the same way, and in different developmental stages distinct mechanisms might be used to establish gene silencing. A surprising example is the discovery that imprinting in extraembryonic cells of the placenta uses a repression mechanism that mainly involves H3K27me3 (Lewis *et al.*, 2004; Umlauf *et al.*, 2004).

Histone methylation at ICRs reflects the heterochromatic H3K9me3 + H4K20me3 modification pattern; however, there are differences in the enzymatic systems and probably also in the recruitment mechanisms. In particular, the nature of the H3K9me3 HMTase in the context of imprinting is still somewhat unclear as Suv39h enzymes do not affect histone methylation at ICRs. Recently, another H3K9-specific HMTase, Setdb1, was found to associate with a particular imprinted region; however, no mechanistic studies have been performed to confirm a function of this enzyme in imprinting (Regha *et al.*, 2007). As in pericentric heterochromatin, H3K9me3 might serve as a binding platform for HP1 proteins, which can then recruit Suv4-20h enzymes to induce H4K20me3 at ICRs. In somatic cells, silencing of imprinted genes strongly depends on DNA methylation. We still do not know to what extent histone modifications contribute to silencing in this context. It could well be possible that histone methylation at ICRs functions downstream of DNA methylation (Henckel *et al.*, 2009). Histone methylation could also represent an additional layer of complexity to ensure stability of the repressed state.

Gene silencing by H3K9 methylation

Average histone modification profiles across genes have revealed a weak correlation between H3K9me3 and gene repression (Figure 1); however, as discussed above, imprinted genes are major targets of H3K9me3. There is also evidence that H3K9me3 is involved in repression of other genes, for example, nuclear receptor targets (Wissmann *et al.*, 2007). Importantly, a lower H3K9 methylation state, H3K9me2, might also have repressive functions. Differentiated cells carry large domains, up to several megabases long, with high levels of H3K9me2, and genes within these domains are repressed (Wen *et al.*, 2009).

In mammals, the different H3K9 methylation states are mediated by several enzymes (Table 2). H3K9me2 is mainly controlled by G9a and the related G9a-like protein (GLP), which function as heterodimers (Tachibana *et al.*, 2005). Suv39h1 and Suv39h2 induce H3K9me3 at heterochromatin. Little is known about the HMTases that are responsible for 'euchromatic' H3K9me3. The only good candidate for such an enzyme is Setdb1, which, as a recombinant enzyme, has poor activity. However, association with the auxiliary factor mAM confers H3K9me3 activity to this enzyme (Wang *et al.*, 2003). Setdb1 is an important functional constituent of the Kap1 corepressor complex that mainly uses H3K9me3 as a means of gene repression (Sripathy *et al.*, 2006). To what extent Setdb1 is really responsible for H3K9me3 *in vivo* remains to be tested.

Different histone methylation states are thought to confer distinct functions. For H3K9 methylation, it is still unclear whether di- or trimethyl states are functionally distinct. The best characterized binding protein, HP1 (Table 1), has affinity to both H3K9me2 and H3K9me3 *in vitro* (Lachner *et al.*, 2001). H3K9 methylation seems to be crucial for HP1 recruitment and binding to heterochromatin, as Suv39h-mutant cells that lose H3K9me3 also lose HP1 from heterochromatin (Lachner *et al.*, 2001). It is also very likely, however, that additional factors can stabilize the binding of HP1 at H3K9me2/3 targets. How silencing is then facilitated by H3K9 methylation is still unclear. The current model suggests that HP1 binding induces a higher grade of chromatin compaction, which would prevent access of transcription factors or RNA Pol II to DNA.

Developmental functions of HMTases

In the previous sections, we discussed functions of histone lysine methylation marks in gene activation and repression, but how important are these mechanisms for normal development? Over the past few years, knockout mice for several HMTases have been established and characterized. In the following sections, we will summarize these data and discuss the functional implications of histone lysine methylation for cell-type identity and regulation of developmental transitions.

Gene activation by H3K4 methylation

Activation of genes often correlates with H3K4me3 at the promoter region. It is not really clear whether H3K4me3 is a consequence of RNA Pol II recruitment or whether this modification represents a poised state for genes that can be easily activated. The major enzymes that induce

H3K4 methylation states in mice are Mll1–5 (mixed lymphoid leukemia) family members, Setd1a/b enzymes and Ash1l (Table 2). Mll proteins are regarded as important regulators of development as homologous proteins in *Drosophila* (trithorax) positively regulate expression of homeotic genes. This function is conserved in mammals and disruption of individual Mll genes in mice leads to reduced expression of Hox genes and developmental defects, as will be outlined below.

Mll1 was the first Mll family member to be functionally analyzed in mice. Dependent on the knockout strategy, Mll1 disruption results in different phenotypes. Truncation of Mll1 in exon 3b or deletion of exons 9–11 both lead to embryonic lethality between E12.5 and E16.5 (Yu et al., 1995; Yagi et al., 1998). Even heterozygous Mll1^{+/-} knockout mice show defects in segment identity that are caused by reduced expression of distinct Hox genes (Hoxa-9 and Hoxc-7).

A very powerful model system to analyze development and differentiation is the hematopoietic system with well-defined stem cells, progenitor populations and differentiated cells. It is noteworthy that Mll1^{-/-} embryos fail to generate or expand hematopoietic stem cells during embryogenesis (Ernst et al., 2004), and, consistent with these data, conditional inactivation of Mll1 in adult mice also disrupts the hematopoietic stem cell compartment (Jude et al., 2007; McMahon et al., 2007). Although the mechanism is not clear, it seems plausible that impaired Hox gene regulation might contribute to these defects in the hematopoietic system (Lawrence et al., 2005).

To understand the function of Mll1-mediated H3K4me₃, another allele was generated that deletes the SET domain region (Terranova et al., 2006). In contrast to all other Mll1 knockout alleles, ΔSET mutants are fully viable and show only slight homeotic transformations. However, in ΔSET mutants, expression of distinct Hox genes is also reduced (Hoxd-4 and Hoxc-8), indicating that the methyltransferase activity of Mll1 is important for gene activation. On the other hand, the relatively mild phenotype of ΔSET and the only modest down-regulation of Hox genes suggest compensatory mechanisms that allow a rather normal development of these mutants.

Mll1 deficiency does not impair overall H3K4 methylation as several other H3K4-specific methyltransferases exist in mice that can compensate for the loss of Mll1 (Table 2). Three other Mll family members have been disrupted in mice (Mll2, Mll3 and Mll5). These mutants show different phenotypes, suggesting different targets or functions. Mll2^{-/-} mice are growth retarded and show embryonic lethality around E11.5 (Glaser et al., 2006). In mutant embryos and ES cell lines, a higher apoptosis rate might be caused by downregulation of the anti-apoptotic factor Bcl2 (Lubitz et al., 2007). In Mll2^{-/-} ES cells, very few genes are misregulated, suggesting that Mll2 has very few unique targets in this cell type (Glaser et al., 2009). One target gene, Magoh2, which is downregulated in Mll2^{-/-} ES cells, loses H3K4me₃ at the promoter region. Interestingly, there is a concomitant increase in H3K27me₃, indicating that chromatin modifications at Mll target genes are negotiated between antagonizing activities of the Mll group HMTases and Ezh enzymes.

Mll3 and Mll4 are present in complexes that share the PTIP (Pax transactivation domain-interacting protein)

subunit, which is not present in Mll1 or Mll2 complexes (Cho et al., 2007). The difference in complex composition might also regulate targeting to distinct sets of genes. Mll3/4 have no apparent effect on Hox gene expression. Instead, they regulate H3K4 methylation at targets of the nuclear hormone receptors, retinoic acid receptor and peroxisome proliferator-activated receptor-γ (Lee et al., 2006, 2008a). Mll3-mutant mice are viable and do not show severe patterning defects; however, differentiation to distinct lineages, for example, adipocytes, is partially impaired, indicating an important function of Mll3 for normal development (Lee et al., 2008a).

Mll5 was regarded as inactive for a long time as no methyltransferase activity of the recombinant protein could be detected. A recent study has shown that Mll5 is a histone H3K4-specific methyltransferase whose activity critically depends on GlcNacetylation of Threonine 440 (Fujiki et al., 2009). Mll5 knockout mice are born at Mendelian ratios; only a few pups die during the first days postpartum. These data indicate that there are no severe developmental defects associated with loss of Mll5; however, closer inspection of the surviving mutants revealed that hematopoietic development is impaired and, in particular, the function of hematopoietic stem cells is reduced (Heuser et al., 2009; Madan et al., 2009; Zhang et al., 2009).

The enzymatic system for H3K4 methylation in mammals is highly complex. Mll proteins are a part of large complexes with common components and distinct interaction partners. Mll1 and Mll2 complexes share the common Menin subunit. Single Mll1 or Mll2 mutants impair H3K4 methylation at only a subset of Hox genes. Disruption of Menin has a much stronger effect and almost all H3K4 methylation is lost from Hox loci (Wang et al., 2009). Much less is known about targets of the other H3K4 HMTases. Although Mll3–5 induce only some H3K4 methylation outside of the Hox loci, Setd1a/Setd1b HMTases have major roles in global H3K4 methylation (Wu et al., 2008). These enzymes are a part of multi-subunit complexes, most similar to the yeast COMPASS complex, and probably have a very general function in transcriptional activation. We expect that knockouts for these enzymes will generate severe defects, maybe even at the cellular level.

Repression of developmental regulators by H3K27 methylation

A classical model system for developmental gene regulation is the regulation of Hox gene expression. In *Drosophila*, the major players that determine transcriptional activity of distinct Hox genes are components of the activating trithorax and the repressive polycomb system, which are primarily recruited to specific DNA sequences (polycomb response elements). The founding member of the PcG family in *Drosophila* is polycomb, a chromo domain-containing protein whose targeting depends on H3K27me₃ by E(z) (Cao et al., 2002). In mammals, many PcG components are conserved; however, the primary targeting mechanism is still unclear, as no polycomb response elements have been identified as yet.

In mice, there are two H3K27-specific HMTases, Ezh1 and Ezh2 (Table 2). Interestingly, their expression

patterns differ as Ezh2 is predominantly expressed in undifferentiated/proliferating cells, whereas Ezh1 is more abundant on terminal differentiation (Margueron *et al.*, 2008; Ezhkova *et al.*, 2009).

H3K27 methylation correlates with gene repression and is enriched at repressed Hox genes and in female mammals at the inactive X chromosome. Thus, deregulation of Ezh enzymes in mice is expected to generate strong phenotypes by affecting the gene activity of important developmental regulators. Indeed, Ezh2 knockout mice show early embryonic lethality around E8.5 (O'Carroll *et al.*, 2001). Ezh2 knockout ES cells can be established and selectively lose H3K27me2 and H3K27me3 while maintaining H3K27me1. Interestingly, at some important developmental genes, H3K27me3 is preserved in Ezh2^{-/-} ES cells (Shen *et al.*, 2008), indicating that Ezh1 and Ezh2 share partially redundant functions.

To analyze the function of H3K27me3 for developmental transitions, conditional Ezh2 knockout mice were established (Su *et al.*, 2003). During B cell development, Ezh2 is highly expressed in early progenitor populations (pro-B and pre-B cells), whereas Ezh1 is only expressed in late stages. Inactivation of Ezh2 in the hematopoietic lineage blocks the transition from pro-B to pre-B cells; however, later stages of B cell development are not impaired (Su *et al.*, 2003). In this model system, transcriptional deregulation has not been investigated and the reason for the developmental block at the pro-B cell stage might be the improper processing of antigen receptor rearrangements.

Ezh2 deletion was also analyzed in skin development, in which basal layer progenitor stages can be distinguished from suprabasal cells that have initiated the program of terminal differentiation. Again, expression of Ezh2 is highest in the progenitor population and decreases with differentiation stage (Ezhkova *et al.*, 2009). Deletion of Ezh2 in developing skin leads to dramatic loss of H3K27me3, resulting in accelerated epidermal differentiation and precocious acquisition of the epidermal barrier. Gene expression analysis of wild-type versus Ezh2-deleted epidermal precursor cells revealed that although H3K27me3 was almost lost in these cells, only approximately 90 genes were deregulated (mostly

upregulated). Most of these genes are normally expressed in the final stages of epidermal development. The premature activation of the terminal differentiation program in the absence of Ezh2 suggests that transcriptional activators for epidermal differentiation are already present in the progenitor stages and that H3K27me3 blocks their access to target promoters (Ezhkova *et al.*, 2009). Most other genes, which also lose H3K27me3 in the absence of Ezh2, are not activated, suggesting that they are repressed by different mechanisms or that appropriate transcriptional activators are not expressed in this tissue. The question as to how H3K27me3 can block the binding of transcription factors has not been addressed. The intriguing possibility that specific binders of H3K27me3, for example, Cbx2 in the context of the PRC1 complex, mediate compaction of the chromatin structure should be tested in future studies.

We think that H3K27me3 has a major role for developmental transitions (Figure 5). In pluripotent cells, developmental regulators are repressed by the bivalent modifications H3K27me3 and H3K4me3. After lineage decision, these genes are either active and carry H3K4me3 or they are inactive and show enriched signals for H3K27me3. Apparently, in committed progenitor cells there are at least two categories of H3K27me3-repressed genes. One category is lineage-appropriate genes that need to be activated in later stages of differentiation. Full activation of these genes depends on removal of H3K27me3, probably by histone demethylases (Agger *et al.*, 2007; Lan *et al.*, 2007) that would then allow binding of transcriptional activators. Lineage-inappropriate genes, the second category, need to be stably repressed by H3K27me3 together with other mechanisms, such as different histone methylation marks or DNA methylation. Aberrant activation of lineage-inappropriate genes might as well be prevented by the lack of transcriptional activators. More detailed analyses are required to distinguish between these possibilities.

Repression of repetitive elements and gene regulation by H3K9 methylation

The first H3K9-specific methyltransferases that were disrupted in mice were the Suv39h1 and Suv39h2

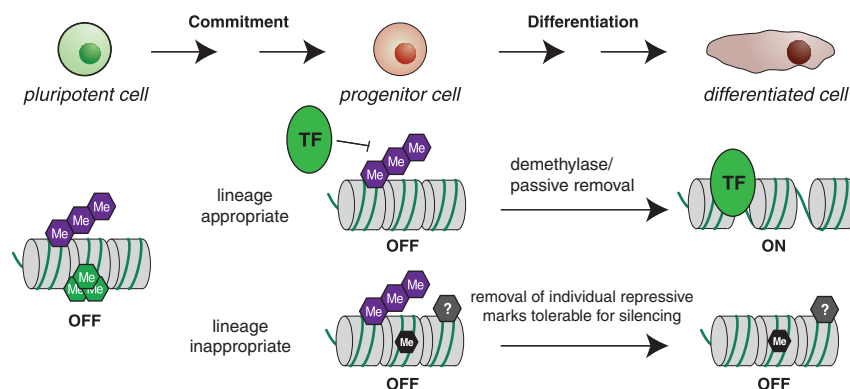


Figure 5 The role of H3K27me3-mediated gene silencing during developmental transitions. In pluripotent cells, developmental regulators are repressed and carry bivalent H3K4me3 + H3K27me3 modifications. Bivalent marks are reduced in committed progenitor cells in which two categories of H3K27me3-repressed genes exist. Lineage-appropriate genes are activated during terminal differentiation, probably by the active removal of H3K27me3. In contrast, lineage-inappropriate genes are stably repressed by H3K27me3 and other mechanisms, such as different histone modifications and DNA methylation.

enzymes (Peters *et al.*, 2001). Single mutants do not show obvious defects, possibly due to compensatory effects. Surprisingly, even Suv39h double mutants are viable, although they show some prenatal lethality during embryogenesis. In Suv39h double-null cells, H3K9me3 is lost from heterochromatin; however, total H3K9me3 levels are only reduced by approximately 50%. Suv39h mainly regulates repetitive regions, such as major satellite repeats and transposons (Martens *et al.*, 2005), as no target genes of Suv39h enzymes could be identified as yet. Alterations in the chromatin structure across repetitive elements can have drastic effects on overall genome stability. It is noteworthy that Suv39h double-null fibroblasts show increased chromosome segregation defects, indicating that centromere function is impaired. Genomic instability could also contribute to tumor development and sterility, which was observed in Suv39h-mutant mice (Peters *et al.*, 2001). These findings show that H3K9me3 has important functions for genomic integrity by repression of mobile elements, thereby ensuring normal development and long-term survival.

Another H3K9-specific HMTase, Setdb1, has been suggested to induce H3K9me3 for gene repression. Setdb1 knockout mice show a strong phenotype with peri-implantation lethality between days E3.5 and E5.5. The defects are so severe that not even embryonic stem cell lines could be established from null mutant blastocysts (Dodge *et al.*, 2004), the reason for which is not understood. Setdb1-mediated H3K9me3 might have important roles in later developmental transitions by regulating targets of transcription factors (Yang *et al.*, 2002) and nuclear hormone receptors (Takada *et al.*, 2007). In a very elegant study, Takada *et al.* (2007) show that in mesenchymal cells of the bone marrow, non-canonical Wnt signaling induces phosphorylation of Setdb1, which leads to the formation of a corepressor complex with peroxisome proliferator-activated receptor- γ . Repression of peroxisome proliferator-activated receptor- γ targets by H3K9me3 drives differentiation of mesenchymal stem cells into the osteoblastic lineage. Catalytically inactive Setdb1 mutants could not ensure repression of peroxisome proliferator-activated receptor- γ targets, leading to differentiation of the mesenchymal cells into adipocytes. These findings show that H3K9me3-mediated gene repression is extremely important to establish transcriptional programs during lineage decisions. Conditional inactivation of Setdb1 in different tissues will surely reveal additional implications of this essential HMTase for normal development.

H3K9me2 is a mainly euchromatic modification that is controlled by G9a and GLP HMTases (Table 2). G9a knockout mice show severe defects during embryogenesis, leading to developmental delay and lethality around day E9.5 (Tachibana *et al.*, 2002). Unlike Suv39h1 and Suv39h2 that work redundantly to induce H3K9me3 at heterochromatin, heterodimer formation between G9a and GLP is essential to establish H3K9me2. Thus, GLP mutant embryos show defects that are almost identical to the G9a knockout, and double inactivation of G9a and GLP does not cause stronger phenotypes (Tachibana *et al.*, 2005).

G9a-mutant ES cells show only little transcriptional deregulation and only one target gene, Mage-2a, has been shown as a direct target (Tachibana *et al.*, 2002). Interestingly, G9a-mediated H3K9 methylation has im-

portant functions during the differentiation of ES cells through the stable inactivation of approximately 120 genes, including the pluripotency genes Oct4 and Nanog (Feldman *et al.*, 2006; Epsztejn-Litman *et al.*, 2008). Stability of silencing is ensured through DNA methylation by Dnmt3a/b, which directly interacts with G9a (Epsztejn-Litman *et al.*, 2008).

G9a mediates H3K9 methylation in another developmentally important context: genomic imprinting. In extraembryonic tissues, imprinting of the Kcnq1 domain is not dependent on DNA methylation, but the ICR shows high levels of H3K9me2 and H3K9me3. It is noteworthy that G9a mutants lose parental-specific imprinting of the Kcnq1 region selectively in the extraembryonic tissue through loss of H3K9me2 and H3K9me3 (Wagschal *et al.*, 2008). Although the average enrichment of H3K9 methylation states across silenced genes (Figure 1) has not suggested major roles for this modification in gene regulation, the abovementioned examples show that both H3K9me2 and H3K9me3 have important functions for gene silencing and for the regulation of developmental transitions.

Concluding remarks

The functional analysis of different HMTases in the context of knockout mouse models has revealed that histone lysine methylation has important roles in facilitating normal development; however, there are still many open questions. Knockout mice for only approximately 50% of mouse HMTases have been generated and analyzed (Table 2), and for many of these proteins we do not even know their enzymatic activity. Several HMTase mutants show early embryonic lethality with pleiotropic defects. A more detailed functional analysis in different tissues is required to better understand their functional implications in developmental processes. The paper by Ezhkova *et al.* (2009) is a very appropriate example for such an analysis.

How are different histone lysine methylation marks interpreted? Are different methylation states really functionally distinct? We only know binding proteins for a subset of positions and methylation states (Table 1). *In vitro*, most binders analyzed so far show only weak affinity to their targets, and the different methylation states cannot be clearly distinguished. Are there other mechanisms *in vivo* that could increase the affinity to their targets?

Epigenetic gene regulation has become a very complex field. In this review, we have only covered the functions of histone lysine methylation in transcriptional regulation; however, there is interplay and dependency between many different epigenetic mechanisms. Chromatin remodelers, for example, recognize histone modifications and can also mediate the exchange of histone variants (Konev *et al.*, 2007). Non-coding RNAs are able to recruit histone-modifying activities that ultimately alter the transcriptional status of targets (Nagano *et al.*, 2008). Histone modifications can be established in sequential pathways in which one modification recruits specific binders and these proteins, in turn, recruit other modifying activities (Schotta *et al.*, 2004). The future challenge in epigenetic research will be to understand this complex network of regulatory mechanisms.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

We apologize to all authors whose work we could not cite because of space limitations. We are grateful to Antonia Jack for a critical reading of the paper. Research in the laboratory of G.S. was financed by the SFB-TR5 Chromatin, Center of Integrated Protein Science Munich and Friedrich Baur-Stiftung.

References

- Agger K, Cloos PA, Christensen J, Pasini D, Rose S, Rappsilber J et al. (2007). UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature* **449**: 731–734.
- Allfrey VG, Faulkner R, Mirsky AE (1964). Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proc Natl Acad Sci USA* **51**: 786–794.
- Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z et al. (2007). High-resolution profiling of histone methylations in the human genome. *Cell* **129**: 823–837.
- Belotserkovskaya R, Reinberg D (2004). Facts about FACT and transcript elongation through chromatin. *Curr Opin Genet Dev* **14**: 139–146.
- Benetti R, Gonzalo S, Jaco I, Schotta G, Klatt P, Jenuwein T et al. (2007). Suv4-20h deficiency results in telomere elongation and derepression of telomere recombination. *J Cell Biol* **178**: 925–936.
- Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J et al. (2006a). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**: 315–326.
- Bernstein E, Duncan EM, Masui O, Gil J, Heard E, Allis CD (2006b). Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. *Mol Cell Biol* **26**: 2560–2569.
- Botuyan MV, Lee J, Ward IM, Kim JE, Thompson JR, Chen J et al. (2006). Structural basis for the methylation state-specific recognition of histone H4-K20 by 53BP1 and Crb2 in DNA repair. *Cell* **127**: 1361–1373.
- Brown MA, Sims III RJ, Gottlieb PD, Tucker PW (2006). Identification and characterization of Smyd2: a split SET/MYND domain-containing histone H3 lysine 36-specific methyltransferase that interacts with the Sin3 histone deacetylase complex. *Mol Cancer* **5**: 26.
- Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P et al. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* **298**: 1039–1043.
- Carrozza MJ, Li B, Florens L, Suganuma T, Swanson SK, Lee KK et al. (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* **123**: 581–592.
- Chamberlain SJ, Yee D, Magnuson T (2008). Polycomb repressive complex 2 is dispensable for maintenance of embryonic stem cell pluripotency. *Stem Cells* **26**: 1496–1505.
- Champagne KS, Saksook N, Pena PV, Johnson K, Ullah M, Yang XJ et al. (2008). The crystal structure of the ING5 PHD finger in complex with an H3K4me3 histone peptide. *Proteins* **72**: 1371–1376.
- Cho YW, Hong T, Hong S, Guo H, Yu H, Kim D et al. (2007). PTIP associates with MLL3- and MLL4-containing histone H3 lysine 4 methyltransferase complex. *J Biol Chem* **282**: 20395–20406.
- Collins RE, Northrop JP, Horton JR, Lee DY, Zhang X, Stallcup MR et al. (2008). The ankyrin repeats of G9a and GLP histone methyltransferases are mono- and dimethyllysine binding modules. *Nat Struct Mol Biol* **15**: 245–250.
- Core LJ, Lis JT (2008). Transcription regulation through promoter-proximal pausing of RNA polymerase II. *Science* **319**: 1791–1792.
- Cui K, Zang C, Roh TY, Schones DE, Childs RW, Peng W et al. (2009). Chromatin signatures in multipotent human hematopoietic stem cells indicate the fate of bivalent genes during differentiation. *Cell Stem Cell* **4**: 80–93.
- Daujat S, Weiss T, Mohn F, Lange UC, Ziegler-Birling C, Zeissler U et al. (2009). H3K64 trimethylation marks heterochromatin and is dynamically remodeled during developmental reprogramming. *Nat Struct Mol Biol* **16**: 777–781.
- Dodge JE, Kang YK, Beppu H, Lei H, Li E (2004). Histone H3-K9 methyltransferase ESET is essential for early development. *Mol Cell Biol* **24**: 2478–2486.
- Edmunds JW, Mahadevan LC, Clayton AL (2008). Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation. *EMBO J* **27**: 406–420.
- Epsztejn-Litman S, Feldman N, Abu-Remaileh M, Shufaro Y, Gerson A, Ueda J et al. (2008). De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. *Nat Struct Mol Biol* **15**: 1176–1183.
- Ernst P, Fisher JK, Avery W, Wade S, Foy D, Korsmeyer SJ (2004). Definitive hematopoiesis requires the mixed-lineage leukemia gene. *Dev Cell* **6**: 437–443.
- Eymery A, Callanan M, Vourc'h C (2009). The secret message of heterochromatin: new insights into the mechanisms and function of centromeric and pericentric repeat sequence transcription. *Int J Dev Biol* **53**: 259–268.
- Ezhkova E, Pasolli HA, Parker JS, Stokes N, Su IH, Hannon G et al. (2009). Ezh2 orchestrates gene expression for the stepwise differentiation of tissue-specific stem cells. *Cell* **136**: 1122–1135.
- Feldman N, Gerson A, Fang J, Li E, Zhang Y, Shinkai Y et al. (2006). G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. *Nat Cell Biol* **8**: 188–194.
- Feng Q, Wang H, Ng HH, Erdjument-Bromage H, Tempst P, Struhl K et al. (2002). Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain. *Curr Biol* **12**: 1052–1058.
- Fischle W, Franz H, Jacobs SA, Allis CD, Khorasanizadeh S (2008). Specificity of the chromodomain Y chromosome family of chromodomains for lysine-methylated ARK(S/T) motifs. *J Biol Chem* **283**: 19626–19635.
- Flanagan JF, Mi LZ, Chruszcz M, Cymborowski M, Clines KL, Kim Y et al. (2005). Double chromodomains cooperate to recognize the methylated histone H3 tail. *Nature* **438**: 1181–1185.
- Fournier C, Goto Y, Ballestar E, Delaval K, Hever AM, Esteller M et al. (2002). Allele-specific histone lysine methylation marks regulatory regions at imprinted mouse genes. *EMBO J* **21**: 6560–6570.
- Fuda NJ, Ardehali MB, Lis JT (2009). Defining mechanisms that regulate RNA polymerase II transcription *in vivo*. *Nature* **461**: 186–192.
- Fujiki R, Chikanishi T, Hashiba W, Ito H, Takada I, Roeder RG et al. (2009). GlcNAcylation of a histone methyltransferase in retinoic-acid-induced granulopoiesis. *Nature* **459**: 455–459.
- Glaser S, Lubitz S, Loveland KL, Ohbo K, Robb L, Schwenk F et al. (2009). The histone 3 lysine 4 methyltransferase, Mll2, is only required briefly in development and spermatogenesis. *Epigenetics Chromatin* **2**: 5.
- Glaser S, Schaft J, Lubitz S, Vintersten K, van der Hoeven F, Tufteland KR et al. (2006). Multiple epigenetic maintenance factors implicated by the loss of Mll2 in mouse development. *Development* **133**: 1423–1432.
- Gonzalo S, Garcia-Cao M, Fraga MF, Schotta G, Peters AH, Cotter SE et al. (2005). Role of the RB1 family in stabilizing histone methylation at constitutive heterochromatin. *Nat Cell Biol* **7**: 420–428.

- Goo YH, Sohn YC, Kim DH, Kim SW, Kang MJ, Jung DJ *et al.* (2003). Activating signal cointegrator 2 belongs to a novel steady-state complex that contains a subset of trithorax group proteins. *Mol Cell Biol* **23**: 140–149.
- Gottlieb PD, Pierce SA, Sims RJ, Yamagishi H, Weihe EK, Harriss JV *et al.* (2002). Bop encodes a muscle-restricted protein containing MYND and SET domains and is essential for cardiac differentiation and morphogenesis. *Nat Genet* **31**: 25–32.
- Gregory GD, Vakoc CR, Rozovskaia T, Zheng X, Patel S, Nakamura T *et al.* (2007). Mammalian ASH1L is a histone methyltransferase that occupies the transcribed region of active genes. *Mol Cell Biol* **27**: 8466–8479.
- Grewal SI, Jia S (2007). Heterochromatin revisited. *Nat Rev Genet* **8**: 35–46.
- Hamamoto R, Furukawa Y, Morita M, Iimura Y, Silva FP, Li M *et al.* (2004). SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells. *Nat Cell Biol* **6**: 731–740.
- Hayashi K, Yoshida K, Matsui Y (2005). A histone H3 methyltransferase controls epigenetic events required for meiotic prophase. *Nature* **438**: 374–378.
- Henckel A, Nakabayashi K, Sanz LA, Feil R, Hata K, Arnaud P (2009). Histone methylation is mechanistically linked to DNA methylation at imprinting control regions in mammals. *Hum Mol Genet* **18**: 3375–3383.
- Heuser M, Yap DB, Leung M, de Algora TR, Tafesh A, McKinney S *et al.* (2009). Loss of MLL5 results in pleiotropic hematopoietic defects, reduced neutrophil immune function, and extreme sensitivity to DNA demethylation. *Blood* **113**: 1432–1443.
- Hoyt PR, Bartholomew C, Davis AJ, Yutzey K, Gamer LW, Potter SS *et al.* (1997). The *Evi1* proto-oncogene is required at midgestation for neural, heart, and paraxial mesenchyme development. *Mech Dev* **65**: 55–70.
- Hung T, Binda O, Champagne KS, Kuo AJ, Johnson K, Chang HY *et al.* (2009). ING4 mediates crosstalk between histone H3 K4 trimethylation and H3 acetylation to attenuate cellular transformation. *Mol Cell* **33**: 248–256.
- Huyen Y, Zgheib O, Ditullio Jr RA, Gorgoulis VG, Zacharatos P, Petty TJ *et al.* (2004). Methylated lysine 79 of histone H3 targets 53BP1 to DNA double-strand breaks. *Nature* **432**: 406–411.
- Jacobson RH, Ladurner AG, King DS, Tjian R (2000). Structure and function of a human TAFII250 double bromodomain module. *Science* **288**: 1422–1425.
- Jamai A, Puglisi A, Strubin M (2009). Histone chaperone spt16 promotes redeposition of the original h3-h4 histones evicted by elongating RNA polymerase. *Mol Cell* **35**: 377–383.
- Jones B, Su H, Bhat A, Lei H, Bajko J, Hevi S *et al.* (2008). The histone H3K79 methyltransferase Dot1L is essential for mammalian development and heterochromatin structure. *PLoS Genet* **4**: e1000190.
- Jude CD, Climer L, Xu D, Artinger E, Fisher JK, Ernst P (2007). Unique and independent roles for MLL in adult hematopoietic stem cells and progenitors. *Cell Stem Cell* **1**: 324–337.
- Kim KC, Geng L, Huang S (2003). Inactivation of a histone methyltransferase by mutations in human cancers. *Cancer Res* **63**: 7619–7623.
- Kim T, Buratowski S (2009). Dimethylation of H3K4 by Set1 recruits the Set3 histone deacetylase complex to 5' transcribed regions. *Cell* **137**: 259–272.
- Kolasinska-Zwierz P, Down T, Latorre I, Liu T, Liu XS, Ahringer J (2009). Differential chromatin marking of introns and expressed exons by H3K36me3. *Nat Genet* **41**: 376–381.
- Konev AY, Tribus M, Park SY, Podhraski V, Lim CY, Emelyanov AV *et al.* (2007). CHD1 motor protein is required for deposition of histone variant H3.3 into chromatin *in vivo*. *Science* **317**: 1087–1090.
- Ku M, Koche RP, Rheinbay E, Mendenhall EM, Endoh M, Mikkelsen TS *et al.* (2008). Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PLoS Genet* **4**: e1000242.
- Kurash JK, Lei H, Shen Q, Marston WL, Granda BW, Fan H *et al.* (2008). Methylation of p53 by Set7/9 mediates p53 acetylation and activity *in vivo*. *Mol Cell* **29**: 392–400.
- Lachner M, O'Carroll D, Rea S, Mechtler K, Jenuwein T (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**: 116–120.
- Lan F, Bayliss PE, Rinn JL, Whetstone JR, Wang JK, Chen S *et al.* (2007). A histone H3 lysine 27 demethylase regulates animal posterior development. *Nature* **449**: 689–694.
- Lawrence HJ, Christensen J, Fong S, Hu YL, Weissman I, Sauvageau G *et al.* (2005). Loss of expression of the Hoxa-9 homeobox gene impairs the proliferation and repopulating ability of hematopoietic stem cells. *Blood* **106**: 3988–3994.
- Lee BM, Mahadevan LC (2009). Stability of histone modifications across mammalian genomes: implications for 'epigenetic' marking. *J Cell Biochem* **108**: 22–34.
- Lee J, Saha PK, Yang QH, Lee S, Park JY, Suh Y *et al.* (2008a). Targeted inactivation of MLL3 histone H3-Lys-4 methyltransferase activity in the mouse reveals vital roles for MLL3 in adipogenesis. *Proc Natl Acad Sci USA* **105**: 19229–19234.
- Lee J, Thompson JR, Botuyan MV, Mer G (2008b). Distinct binding modes specify the recognition of methylated histones H3K4 and H4K20 by JMJD2A-tudor. *Nat Struct Mol Biol* **15**: 109–111.
- Lee JH, Tate CM, You JS, Skalknik DG (2007). Identification and characterization of the human Set1B histone H3-Lys4 methyltransferase complex. *J Biol Chem* **282**: 13419–13428.
- Lee S, Lee DK, Dou Y, Lee J, Lee B, Kwak E *et al.* (2006). Coactivator as a target gene specificity determinant for histone H3 lysine 4 methyltransferases. *Proc Natl Acad Sci USA* **103**: 15392–15397.
- Lee SH, Oshige M, Durant ST, Rasila KK, Williamson EA, Ramsey H *et al.* (2005). The SET domain protein Metnase mediates foreign DNA integration and links integration to nonhomologous end-joining repair. *Proc Natl Acad Sci USA* **102**: 18075–18080.
- Leeb M, Pasini D, Novatchkova M, Jaritz M, Helin K, Wutz A (2010). Polycomb complexes act redundantly to repress genomic repeats and genes. *Genes Dev* **24**: 265–276.
- Lewis A, Mitsuya K, Umlauf D, Smith P, Dean W, Walter J *et al.* (2004). Imprinting on distal chromosome 7 in the placenta involves repressive histone methylation independent of DNA methylation. *Nat Genet* **36**: 1291–1295.
- Liu H, Westergaard TD, Hsieh JJ (2009). MLL5 governs hematopoiesis: a step closer. *Blood* **113**: 1395–1396.
- Loyola A, Tagami H, Bonaldi T, Roche D, Quivy JP, Imhof A *et al.* (2009). The HP1alpha-CAF1-SetDB1-containing complex provides H3K9me1 for Suv39-mediated K9me3 in pericentric heterochromatin. *EMBO Rep* **10**: 769–775.
- Lubitiz S, Glaser S, Schaft J, Stewart AF, Anastassiadis K (2007). Increased apoptosis and skewed differentiation in mouse embryonic stem cells lacking the histone methyltransferase Mll2. *Mol Biol Cell* **18**: 2356–2366.
- Madan V, Madan B, Brykczynska U, Zilbermann F, Hogeveen K, Dohner K *et al.* (2009). Impaired function of primitive hematopoietic cells in mice lacking the Mixed-Lineage-Leukemia homolog MLL5. *Blood* **113**: 1444–1454.
- Margueron R, Justin N, Ohno K, Sharpe ML, Son J, Drury III WJ *et al.* (2009). Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature* **461**: 762–767.
- Margueron R, Li G, Sarma K, Blais A, Zavadil J, Woodcock CL *et al.* (2008). Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. *Mol Cell* **32**: 503–518.
- Martens JH, O'Sullivan RJ, Braunschweig U, Opravil S, Radolf M, Steinlein P *et al.* (2005). The profile of repeat-associated

- histone lysine methylation states in the mouse epigenome. *EMBOJ* **24**: 800–812.
- Matthews AG, Kuo AJ, Ramon-Maiques S, Han S, Champagne KS, Ivanov D *et al.* (2007). RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. *Nature* **450**: 1106–1110.
- McMahon KA, Hiew SY, Hadjur S, Veiga-Fernandes H, Menzel U, Price AJ *et al.* (2007). Mll has a critical role in fetal and adult hematopoietic stem cell self-renewal. *Cell Stem Cell* **1**: 338–345.
- Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G *et al.* (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**: 553–560.
- Milne TA, Briggs SD, Brock HW, Martin ME, Gibbs D, Allis CD *et al.* (2002). MLL targets SET domain methyltransferase activity to Hox gene promoters. *Mol Cell* **10**: 1107–1117.
- Milne TA, Dou Y, Martin ME, Brock HW, Roeder RG, Hess JL (2005). MLL associates specifically with a subset of transcriptionally active target genes. *Proc Natl Acad Sci USA* **102**: 14765–14770.
- Nagano T, Mitchell JA, Sanz LA, Pauler FM, Ferguson-Smith AC, Feil R *et al.* (2008). The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science* **322**: 1717–1720.
- Nimura K, Ura K, Shiratori H, Ikawa M, Okabe M, Schwartz RJ *et al.* (2009). A histone H3 lysine 36 trimethyltransferase links Nkx2-5 to Wolf-Hirschhorn syndrome. *Nature* **460**: 287–291.
- Nishioka K, Chuikov S, Sarma K, Erdjument-Bromage H, Allis CD, Tempst P *et al.* (2002). Set9, a novel histone H3 methyltransferase that facilitates transcription by precluding histone tail modifications required for heterochromatin formation. *Genes Dev* **16**: 479–489.
- O'Carroll D, Erhardt S, Pagani M, Barton SC, Surani MA, Jenuwein T (2001). The polycomb-group gene *Ezh2* is required for early mouse development. *Mol Cell Biol* **21**: 4330–4336.
- Oda H, Okamoto I, Murphy N, Chu J, Price SM, Shen MM *et al.* (2009). Monomethylation of histone H4-lysine 20 is involved in chromosome structure and stability and is essential for mouse development. *Mol Cell Biol* **29**: 2278–2295.
- Ohinata Y, Payer B, O'Carroll D, Ancelin K, Ono Y, Sano M *et al.* (2005). Blimp1 is a critical determinant of the germ cell lineage in mice. *Nature* **436**: 207–213.
- Pan G, Tian S, Nie J, Yang C, Ruotti V, Wei H *et al.* (2007). Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. *Cell Stem Cell* **1**: 299–312.
- Pannetier M, Julien E, Schotta G, Tardat M, Sardet C, Jenuwein T *et al.* (2008). PR-SET7 and SUV4-20H regulate H4 lysine-20 methylation at imprinting control regions in the mouse. *EMBO Rep* **9**: 998–1005.
- Pauler FM, Sloane MA, Huang R, Regha K, Koerner MV, Tamir I *et al.* (2009). H3K27me3 forms BLOCs over silent genes and intergenic regions and specifies a histone banding pattern on a mouse autosomal chromosome. *Genome Res* **19**: 221–233.
- Pavri R, Zhu B, Li G, Trojer P, Mandal S, Shilatifard A *et al.* (2006). Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. *Cell* **125**: 703–717.
- Peters AH, O'Carroll D, Scherthan H, Mechtler K, Sauer S, Schofer C *et al.* (2001). Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* **107**: 323–337.
- Rayasam GV, Wendling O, Angrand PO, Mark M, Niederreither K, Song L *et al.* (2003). NSD1 is essential for early post-implantation development and has a catalytically active SET domain. *EMBOJ* **22**: 3153–3163.
- Rea S, Eisenhaber F, O'Carroll D, Strahl BD, Sun ZW, Schmid M *et al.* (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406**: 593–599.
- Regha K, Sloane MA, Huang R, Pauler FM, Warczok KE, Melikant B *et al.* (2007). Active and repressive chromatin are interspersed without spreading in an imprinted gene cluster in the mammalian genome. *Mol Cell* **27**: 353–366.
- Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Bruggmann SA *et al.* (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **129**: 1311–1323.
- Schoeftner S, Blasco MA (2008). Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. *Nat Cell Biol* **10**: 228–236.
- Schoeftner S, Sengupta AK, Kubicek S, Mechtler K, Spahn L, Koseki H *et al.* (2006). Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing. *EMBOJ* **25**: 3110–3122.
- Schotta G, Lachner M, Sarma K, Ebert A, Sengupta R, Reuter G *et al.* (2004). A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev* **18**: 1251–1262.
- Schotta G, Sengupta R, Kubicek S, Malin S, Kauer M, Callen E *et al.* (2008). A chromatin-wide transition to H4K20 mono-methylation impairs genome integrity and programmed DNA rearrangements in the mouse. *Genes Dev* **22**: 2048–2061.
- Shen X, Liu Y, Hsu YJ, Fujiwara Y, Kim J, Mao X *et al.* (2008). EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. *Mol Cell* **32**: 491–502.
- Shi X, Hong T, Walter KL, Ewalt M, Michishita E, Hung T *et al.* (2006). ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature* **442**: 96–99.
- Spies N, Nielsen CB, Padgett RA, Burge CB (2009). Biased chromatin signatures around polyadenylation sites and exons. *Mol Cell* **36**: 245–254.
- Sripathy SP, Stevens J, Schultz DC (2006). The KAP1 corepressor functions to coordinate the assembly of *de novo* HP1-demarcated microenvironments of heterochromatin required for KRAB zinc finger protein-mediated transcriptional repression. *Mol Cell Biol* **26**: 8623–8638.
- Steele-Perkins G, Fang W, Yang XH, Van Gele M, Carling T, Gu J *et al.* (2001). Tumor formation and inactivation of RIZ1, an Rb-binding member of a nuclear protein-methyltransferase superfamily. *Genes Dev* **15**: 2250–2262.
- Strahl BD, Ohba R, Cook RG, Allis CD (1999). Methylation of histone H3 at lysine 4 is highly conserved and correlates with transcriptionally active nuclei in Tetrahymena. *Proc Natl Acad Sci USA* **96**: 14967–14972.
- Su IH, Basavaraj A, Krutchinsky AN, Hobert O, Ullrich A, Chait BT *et al.* (2003). Ezh2 controls B cell development through histone H3 methylation and IgH rearrangement. *Nat Immunol* **4**: 124–131.
- Sun XJ, Wei J, Wu XY, Hu M, Wang L, Wang HH *et al.* (2005). Identification and characterization of a novel human histone H3 lysine 36-specific methyltransferase. *J Biol Chem* **280**: 35261–35271.
- Sun Y, Jiang X, Xu Y, Ayrapetov MK, Moreau LA, Whetstone JR *et al.* (2009). Histone H3 methylation links DNA damage detection to activation of the tumour suppressor Tip60. *Nat Cell Biol* **11**: 1376–1382.
- Tachibana M, Matsumura Y, Fukuda M, Kimura H, Shinkai Y (2008). G9a/GLP complexes independently mediate H3K9 and DNA methylation to silence transcription. *EMBOJ* **27**: 2681–2690.
- Tachibana M, Sugimoto K, Nozaki M, Ueda J, Ohta T, Ohki M *et al.* (2002). G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes Dev* **16**: 1779–1791.
- Tachibana M, Ueda J, Fukuda M, Takeda N, Ohta T, Iwanari H *et al.* (2005). Histone methyltransferases G9a and GLP form

- heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. *Genes Dev* **19**: 815–826.
- Takada I, Mihara M, Suzawa M, Ohtake F, Kobayashi S, Igarashi M et al. (2007). A histone lysine methyltransferase activated by non-canonical Wnt signalling suppresses PPAR- γ transactivation. *Nat Cell Biol* **9**: 1273–1285.
- Tan X, Rotllant J, Li H, De Deyne P, Du SJ (2006). SmyD1, a histone methyltransferase, is required for myofibril organization and muscle contraction in zebrafish embryos. *Proc Natl Acad Sci USA* **103**: 2713–2718.
- Taverna SD, Ilin S, Rogers RS, Tanny JC, Lavender H, Li H et al. (2006). Yng1 PHD finger binding to H3 trimethylated at K4 promotes NuA3 HAT activity at K14 of H3 and transcription at a subset of targeted ORFs. *Mol Cell* **24**: 785–796.
- Terranova R, Agherbi H, Boned A, Meresse S, Djabali M (2006). Histone and DNA methylation defects at Hox genes in mice expressing a SET domain-truncated form of Mll. *Proc Natl Acad Sci USA* **103**: 6629–6634.
- Trojer P, Li G, Sims III RJ, Vaquero A, Kalakonda N, Bocconi P et al. (2007). L3MBTL1, a histone-methylation-dependent chromatin lock. *Cell* **129**: 915–928.
- Umlauf D, Goto Y, Cao R, Cerqueira F, Wagschal A, Zhang Y et al. (2004). Imprinting along the Kcnq1 domain on mouse chromosome 7 involves repressive histone methylation and recruitment of Polycomb group complexes. *Nat Genet* **36**: 1296–1300.
- Vermeulen M, Mulder KW, Denissov S, Pijnappel WW, van Schaik FM, Varier RA et al. (2007). Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* **131**: 58–69.
- Vire E, Brenner C, Deplus R, Blanchon L, Fraga M, Didelot C et al. (2006). The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* **439**: 871–874.
- Wagschal A, Sutherland HG, Woodfine K, Henckel A, Chebli K, Schulz R et al. (2008). G9a histone methyltransferase contributes to imprinting in the mouse placenta. *Mol Cell Biol* **28**: 1104–1113.
- Wang H, An W, Cao R, Xia L, Erdjument-Bromage H, Chatton B et al. (2003). mAM facilitates conversion by ESET of dimethyl to trimethyl lysine 9 of histone H3 to cause transcriptional repression. *Mol Cell* **12**: 475–487.
- Wang H, Cao R, Xia L, Erdjument-Bromage H, Borchers C, Tempst P et al. (2001). Purification and functional characterization of a histone H3-lysine 4-specific methyltransferase. *Mol Cell* **8**: 1207–1217.
- Wang H, Wang L, Erdjument-Bromage H, Vidal M, Tempst P, Jones RS et al. (2004). Role of histone H2A ubiquitination in Polycomb silencing. *Nature* **431**: 873–878.
- Wang P, Lin C, Smith ER, Guo H, Sanderson BW, Wu M et al. (2009). Global analysis of H3K4 methylation defines MLL family member targets and points to a role for MLL1-mediated H3K4 methylation in the regulation of transcriptional initiation by RNA polymerase II. *Mol Cell Biol* **29**: 6074–6085.
- Wen B, Wu H, Shinkai Y, Irizarry RA, Feinberg AP (2009). Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nat Genet* **41**: 246–250.
- Wissmann M, Yin N, Muller JM, Greschik H, Fodor BD, Jenuwein T et al. (2007). Cooperative demethylation by JMJD2C and LSD1 promotes androgen receptor-dependent gene expression. *Nat Cell Biol* **9**: 347–353.
- Wu M, Wang PF, Lee JS, Martin-Brown S, Florens L, Washburn M et al. (2008). Molecular regulation of H3K4 trimethylation by Wdr82, a component of human Set1/COMPASS. *Mol Cell Biol* **28**: 7337–7344.
- Wysocka J, Myers MP, Laherty CD, Eisenman RN, Herr W (2003). Human Sin3 deacetylase and trithorax-related Set1/Ash2 histone H3-K4 methyltransferase are tethered together selectively by the cell-proliferation factor HCF-1. *Genes Dev* **17**: 896–911.
- Wysocka J, Swigut T, Xiao H, Milne TA, Kwon SY, Landry J et al. (2006). A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* **442**: 86–90.
- Yagi H, Deguchi K, Aono A, Tani Y, Kishimoto T, Komori T (1998). Growth disturbance in fetal liver hematopoiesis of Mll-mutant mice. *Blood* **92**: 108–117.
- Yamaji M, Seki Y, Kurimoto K, Yabuta Y, Yuasa M, Shigeta M et al. (2008). Critical function of Prdm14 for the establishment of the germ cell lineage in mice. *Nat Genet* **40**: 1016–1022.
- Yang L, Xia L, Wu DY, Wang H, Chansky HA, Schubach WH et al. (2002). Molecular cloning of ESET, a novel histone H3-specific methyltransferase that interacts with ERG transcription factor. *Oncogene* **21**: 148–152.
- Yochum GS, Ayer DE (2002). Role for the mortality factors MORF4, MRGX, and MRG15 in transcriptional repression via associations with Pf1, mSin3A, and Transducin-Like Enhancer of Split. *Mol Cell Biol* **22**: 7868–7876.
- Yoh SM, Lucas JS, Jones KA (2008). The Iws1:Spt6:CTD complex controls cotranscriptional mRNA biosynthesis and HYPB/Setd2-mediated histone H3K36 methylation. *Genes Dev* **22**: 3422–3434.
- Yu BD, Hess JL, Horning SE, Brown GA, Korsmeyer SJ (1995). Altered Hox expression and segmental identity in Mll-mutant mice. *Nature* **378**: 505–508.
- Zhang P, Du J, Sun B, Dong X, Xu G, Zhou J et al. (2006). Structure of human MRG15 chromo domain and its binding to Lys36-methylated histone H3. *Nucleic Acids Res* **34**: 6621–6628.
- Zhang Y, Wong J, Klinger M, Tran MT, Shannon KM, Killeen N (2009). MLL5 contributes to hematopoietic stem cell fitness and homeostasis. *Blood* **113**: 1455–1463.
- Zhou W, Zhu P, Wang J, Pascual G, Ohgi KA, Lozach J et al. (2008). Histone H2A monoubiquitination represses transcription by inhibiting RNA polymerase II transcriptional elongation. *Mol Cell* **29**: 69–80.

2.3 Suv4-20h histone methyltransferases promote neuroectodermal differentiation by silencing the pluripotency-associated Oct-25 gene.

Article: Nicetto et al. 2013, PLOS Genetics

Suv4-20h Histone Methyltransferases Promote Neuroectodermal Differentiation by Silencing the Pluripotency-Associated Oct-25 Gene

Dario Nicetto¹, Matthias Hahn², Julia Jung³, Tobias D. Schneider¹, Tobias Straub¹, Robert David^{3¶}, Gunnar Schotta², Ralph A. W. Rupp^{1*}

1 Adolf Butenandt Institut, Institut für Molekularbiologie, Ludwig Maximilians-Universität, München, Germany, **2** Center for Integrated Protein Science (Munich) at the Institut für Molekularbiologie, Adolf-Butenandt-Institut, LMU, München, Germany, **3** Medizinische Klinik I am Klinikum der Universität München (LMU), München, Germany

Abstract

Post-translational modifications (PTMs) of histones exert fundamental roles in regulating gene expression. During development, groups of PTMs are constrained by unknown mechanisms into combinatorial patterns, which facilitate transitions from uncommitted embryonic cells into differentiated somatic cell lineages. Repressive histone modifications such as H3K9me3 or H3K27me3 have been investigated in detail, but the role of H4K20me3 in development is currently unknown. Here we show that *Xenopus laevis* Suv4-20h1 and h2 histone methyltransferases (HMTases) are essential for induction and differentiation of the neuroectoderm. Morpholino-mediated knockdown of the two HMTases leads to a selective and specific downregulation of genes controlling neural induction, thereby effectively blocking differentiation of the neuroectoderm. Global transcriptome analysis supports the notion that these effects arise from the transcriptional deregulation of specific genes rather than widespread, pleiotropic effects. Interestingly, morphant embryos fail to repress the Oct4-related *Xenopus* gene Oct-25. We validate Oct-25 as a direct target of xSu4-20h enzyme mediated gene repression, showing by chromatin immunoprecipitation that it is decorated with the H4K20me3 mark downstream of the promoter in normal, but not in double-morphant, embryos. Since knockdown of Oct-25 protein significantly rescues the neural differentiation defect in xSuv4-20h double-morphant embryos, we conclude that the epistatic relationship between Suv4-20h enzymes and Oct-25 controls the transit from pluripotent to differentiation-competent neural cells. Consistent with these results in *Xenopus*, murine Suv4-20h1/h2 double-knockout embryonic stem (DKO ES) cells exhibit increased Oct4 protein levels before and during EB formation, and reveal a compromised and biased capacity for *in vitro* differentiation, when compared to normal ES cells. Together, these results suggest a regulatory mechanism, conserved between amphibians and mammals, in which H4K20me3-dependent restriction of specific POU-V genes directs cell fate decisions, when embryonic cells exit the pluripotent state.

Citation: Nicetto D, Hahn M, Jung J, Schneider TD, Straub T, et al. (2013) Suv4-20h Histone Methyltransferases Promote Neuroectodermal Differentiation by Silencing the Pluripotency-Associated Oct-25 Gene. PLoS Genet 9(1): e1003188. doi:10.1371/journal.pgen.1003188

Editor: Anne C. Ferguson-Smith, University of Cambridge, United Kingdom

Received: April 3, 2012; **Accepted:** November 7, 2012; **Published:** January 31, 2013

Copyright: © 2013 Nicetto et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: RAWP was supported by SFB/TR5-DFG (<http://www.dfg.de/foerderung/programme/listen/projektdetails/index.jsp?id=5485239>). GS was supported by SFB/TR5-DFG (<http://www.dfg.de/foerderung/programme/listen/projektdetails/index.jsp?id=5485239>). This work was also supported by Episy-BMBF (<http://www.episys.org/EPISYS/Home.html>). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: ralph.rupp@med.uni-muenchen.de

¶ Current address: Biomedical Research Center Rostock, Rostock University, Rostock, Germany

Introduction

Embryonic development is controlled by fine-tuned differential gene expression. A succession of regulatory protein networks unfolds the zygotic gene expression program along a hierarchy of decisions, leading from the embryonic ground state to the epiblast and then to germ layers, which become patterned into cell type and organ precursor territories. The pluripotent trait, key feature of embryonic stem (ES) cells [1], is progressively restricted and finally lost as soon as embryonic cells become specified to germ layer fates. Recent studies have revealed that alterations in chromatin structure, dynamics and composition represent fundamental processes, which define the epigenetic landscape that directs cell type specification along this hierarchy [2,3].

Besides important contributions from ATP dependent chromatin remodelling factors [4,5] and histone variants [6] in modulating nucleosome dynamics, histone post-translational modifications (PTMs) have been linked to gene expression [3,7]. The transition from pluripotent to differentiated cells is characterized by a progressive increase in heterochromatin formation, in a process that changes the hyperdynamic open chromatin structure into a less accessible architecture [1,8]. At the same time transcriptional silencing of non-lineage specific genes is achieved via acquisition of repressive histone marks. *In vivo* studies have shown that dynamic alterations in the levels of histone modifications characterize early stages of development both in mammals [9–11] and other vertebrates [12–14].

Lysine methylation of histones is catalyzed by SET domain-containing histone methyltransferases (HMTases), and can be

Author Summary

The quest of modern developmental biology is a detailed molecular description of the process that leads from the fertilized egg to the complex and highly differentiated adult organism. This process is controlled largely on the level of gene expression. While early embryonic cells are pluripotent and capable of transcribing most of their genome, older cells have become committed to the germ layer and differentiation programs during gastrulation. They express then a subset of genes compatible with their future physiological function. Young, pluripotent cells and post-gastrula, committed cells express different networks of transcription factors and contain chromatin of different structure and composition. How these two regulatory layers are interconnected during development is incompletely understood. We describe a novel and unexpected link between the pluripotency-associated POU-V gene Oct-25 and xSuv4-20h histone methyltransferases. Xsuv4-20h enzymes are required to repress the Oct-25 gene, a homolog of mammalian Oct4, in the neuroectoderm of frog embryos as a prerequisite for neural differentiation. Consistently, murine Suv4-20h double-null ES cells show increased Oct4 protein levels before and during *in vitro* differentiation and display compromised differentiation in comparison to wild-type ES cells. Thus, Suv4-20h enzymes control specific POU-V genes and are involved in germ-layer specific differentiation.

linked both to transcriptional activation and repression [15,16]. In particular, repressive histone methyl marks are found on lysine residues at position 9 and 27 on histone H3 and in position 20 on histone H4. H3K27 trimethylation is catalyzed by polycomb repression complex (PRC) 2, which predominantly represses developmental regulatory genes [17–19]. Trimethylation of H3K9 and H4K20 relies on Suv39h and Suv4-20h enzyme activities, respectively [20,21], and predominantly marks repetitive genomic DNA at pericentromeric and telomeric heterochromatin [16,21]. While H3K9-specific HMTases have been characterized in significant depth [20,22,23], we know little about the functions of Suv4-20h1 and Suv4-20h2 enzymes with regard to gene regulation. *In vivo* analysis of H4K20 methylation states in mouse embryos reveals specific patterns both in cellular or subnuclear abundance [9,24]. Suv4-20h DKO pups die perinatally, indicating an essential function of the two enzymes during embryogenesis [24]. Moreover, quantitative analysis of histone PTMs in *X. laevis* revealed a progressive and significant accumulation of H4K20me3 levels during embryogenesis, suggesting developmental functions for these enzymes [13].

To characterize the functional role of H4K20me2/3 during vertebrate development we have investigated the consequences of both morpholino-mediated protein knockdown and mRNA-born overexpression of the *Xenopus* Suv4-20h1 and h2 homologs in frog embryos. Our data reveal a specific and selective requirement for Suv4-20h enzyme activities in neuroectodermal differentiation, in a process which involves transcriptional repression of pluripotency associated POU-V genes, both in *Xenopus* embryos and in murine ES cells.

Results

Characterization of *Xenopus* Suv4-20h1 and h2 enzymes

We initially identified *X. laevis* Suv4-20h1 and h2 ESTs via database mining. Mouse and frog Suv4-20h1 and h2 protein

sequences are well conserved, particularly within the SET domains ($\geq 88\%$ identity), even though the xSuv4-20h2 open reading frame is longer than its mouse homolog due to C-terminal insertions (supplementary data, Figure S1). XSuv4-20h1/h2 genes are both maternally and zygotically expressed in a ubiquitous manner, as shown by RNA *in situ* hybridisation and RT-PCR analysis (Figure S2A–S2D). XSuv4-20h1 mRNA abundance decreases during the initial stages of development and subsequently rises from mid-gastrula on, reflecting the switch from maternal-to-zygotic transcription at midblastula. In contrast, the initially high xSuv4-20h2 mRNA level falls and stays low at late stages (Figure S2D).

To test the activities of these *Xenopus* HMTases, we first analyzed their ability to rescue H4K20me3 levels in Suv4-20h1/h2 DKO mouse embryonic fibroblasts (MEF Suv4-20h DKO; [24]), which lack this modification. Both frog cDNAs re-established a proper H4K20me3 pattern, which was strongly enriched at heterochromatic regions that were identified as DAPI-dense chromocenters within nuclei (Figure 1A). Thus, *Xenopus laevis* Suv4-20h homologs are biologically active and can direct H4K20 trimethylation. To test, whether they generate this histone mark in frog embryos, we designed antisense Morpholino oligonucleotides (MO) to reduce synthesis of xSuv4-20h1 and h2 proteins from endogenous mRNAs (Figure S3A, S3B). These MOs specifically inhibited translation of their cognate templates *in vitro* (Figure S3C). To avoid possible functional complementation between the xSuv4-20h enzymes *in vivo*, we decided to inject the two MOs simultaneously into both blastomeres of 2-cell stage embryos and performed western blots with nuclear protein extracts from these double-morphant embryos at the tadpole stage (NF30–33). Compared to uninjected controls or embryos injected with an unrelated control MO (control-morphants), the double morphants contained significantly less H4K20me2 ($p = 0.0011$) and H4K20me3 ($p = 0.0164$), which was coupled to an increase in H4K20me1 ($p = 0.0034$) (Figure 1B and 1C).

This result was confirmed by MALDI-TOF mass spectrometry (Figure S4A). As described in Schneider et al. [13], the relative abundance of histone modifications was calculated by quantifying the amount of a specific modification relative to the amount of all modification states determined for the same histone peptide. As reported before [13], the H4K20me3 mark could not be quantitated reproducibly for technical reasons. Compared to control embryos, however, xSuv4-20h double morphants contained approximately 2.5-fold less of H4K20me2 ($p = 0.0153$) and three-fold more H4K20me1 ($p = 0.0185$), while the abundance of the unmodified peptide state remained unaffected. Importantly, the levels of histone H3 methylation on two tryptic peptides, covering the K9, K27 and K36 positions, were indistinguishable between control and double-morphant embryos (Figure S4B and S4C). Western blot analysis with antibodies against trimethylated H3K9 or H3K27 also showed no difference in the abundance of these two marks between control embryos and xSuv4-20h double morphants (Figure S4D).

To further characterize the effects of xSuv4-20h enzyme depletion on the cellular level, we performed immunohistochemistry on sections from tailbud stage embryos (NF30), which were injected with the xSuv4-20h MO-mix into one of two blastomeres at 2-cell stage together with fluorescently labelled dextrans as lineage tracer. While H3 staining was comparable between injected and uninjected sides under all conditions (Figure S5A), staining for H4K20me2 and –me3 was clearly reduced on the double-morphant side of the neural tube (Figure 1D). In

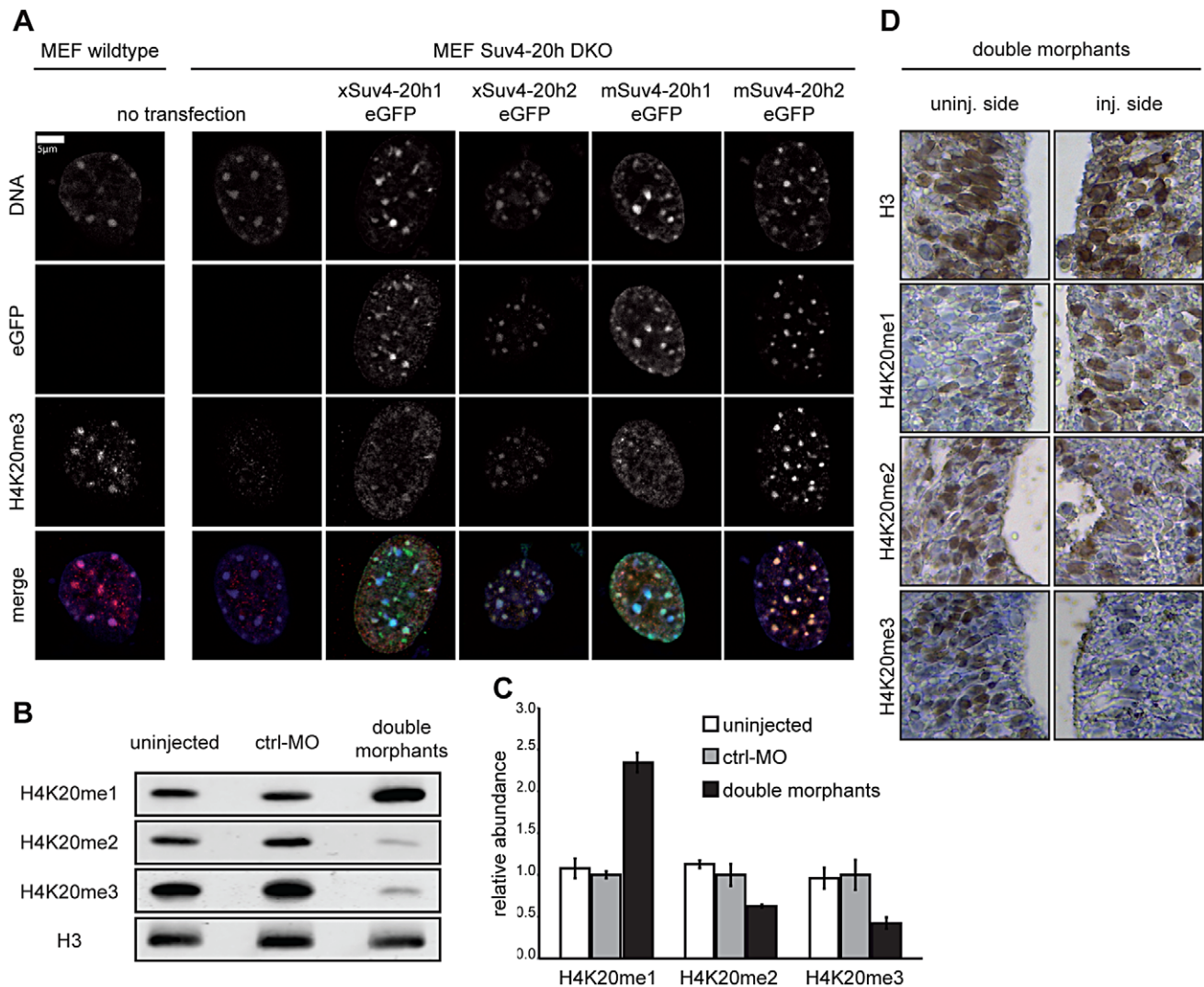


Figure 1. Functional analysis of xSuv4-20h HMTases. (A) Transiently transfected eGFP-tagged Suv4-20h1 and h2 enzymes from frog or mouse re-establish H4K20me3 marks in heterochromatic foci of Suv4-20h1/h2 DKO MEFs. (B–D) Bulk histones from tadpoles (NF30–33) injected with morpholinos targeting translation of endogenous xSuv4-20h1 and h2 mRNA show a strong reduction in H4K20me2 and H4K20me3 levels and a concomitant increase in the H4K20me1 mark. (B) Western Blot analysis of uninjected embryos, control morphants (ctrl-MO), and double morphants with antibodies against H4K20 mono-, di- and trimethylation. PanH3 antibody was used as loading control. (C) Western Blot quantification of three independent biological experiments; data represent mean values, error bars indicate SEM. (D) Immuno-histochemistry on xSuv4-20h double morphant tadpoles. Panels show details from neural tubes stained with antibodies against the histone epitopes indicated on the side. Whole sections shown in Figure S5A.

doi:10.1371/journal.pgen.1003188.g001

agreement with our western blot and mass spec results, the reduction in the di- and tri-methyl mark was coupled to an increase in H4K20me1 staining. Altogether these results indicate that xSuv4-20h1 and h2 downregulation leads to a quantitative reduction of H4K20 di- and trimethyl marks in the frog embryo, without affecting the bulk abundance of other repressive histone marks such as H3 K9/K27 methylation.

RNA-based overexpression of Suv4-20h HMTases had the opposite effect. When injected singly, xSuv4-20h1 or h2 mRNAs increased both di- and trimethylated H4K20 in a dose-dependent manner (Figure S8A). A comparable result was achieved by injection of either mouse Suv4-20h1 or h2 mRNAs (Figure S9A). Together, these results identify the frog cDNAs as orthologs of mammalian Suv4-20h enzymes. Loss and gain of function experiments also indicate that the bulk abundance of di- and

trimethylated H4K20 can be manipulated over a wide range without compromising embryonic viability.

XSuv4-20h HMTases depletion inhibits eye and melanocyte formation

We next tested, whether depletion of xSuv4-20h HMTases affects embryonic development. We injected xSuv4-20h1/h2 MO-mix into one blastomere of two-cell stage embryos and scored phenotypic alterations by comparing injected with uninjected sides. No obvious differences were observed during early development, including gastrulation, axial extension and dorso-ventral patterning. From tailbud stages on, two main phenotypes became manifest. First, in the injected side of xSuv4-20h double morphants the eye formation was strongly compromised. The eye rudiments contained no or barely visible retinal pigment and

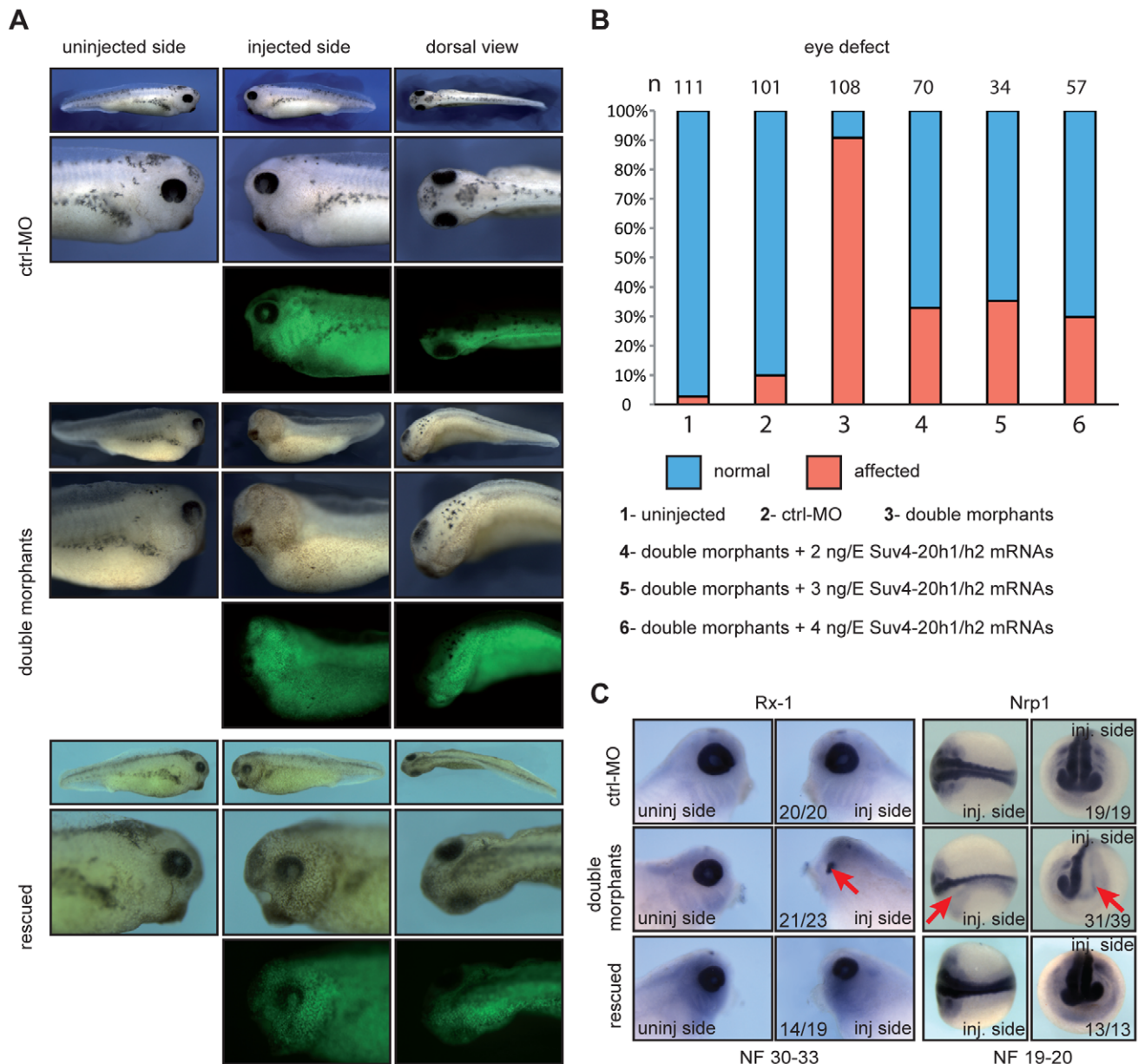


Figure 2. xSuv4-20h1/h2 double morphants lack eyes and melanophores. (A) Morphological phenotypes of representative tadpoles (NF30-33) from embryo cohorts injected into one of two blastomeres at two-cell stage with ctrl-MO, xSuv4-20h1/h2 MOs (double morphants), and double morphants coinjected with 3 ng mouse Suv4-20h1/h2 mRNAs (rescued). Injected body halves were identified by green fluorescence of the coinjected lineage tracer Fluor 488 Dextran. (B) Penetrance of the eye phenotype. Data from three to five independent experiments; n = total number of embryos scored. (C) RNA *in situ* hybridization analysis for Rx-1 in tadpoles (NF30-33), and CNS markers Nrp1 at neural tube stage (NF19-20). For each condition, numbers refer to embryos showing the displayed morphology or staining, in comparison to the total number of analysed embryos. doi:10.1371/journal.pgen.1003188.g002

typically had no lens (Figure 2A). Secondly, melanophores that are found on the dorsal part of the head and the lateral portion of the trunk, were severely reduced in numbers or completely lost from the double-morphant side (Figure 2A). Both phenotypes had a penetrance between 80–90% in xSuv4-20h double morphants ($p < 0.0001$, Fisher's exact test) in several independent experiments (Figure 2B). Control-morphant embryos had normal eyes and melanocytes (Figure 2A) and were indistinguishable from uninjected siblings in most cases (Figure 2B).

The distinct eye phenotype prompted us to investigate the underlying molecular changes. RNA *in situ* hybridization experiments revealed a clearly reduced expression of the homeobox

transcription factor Rx-1 (Figure 2C) and the paired box transcription factor Pax-6 (Figure S5D) in xSuv4-20h double morphants. The reduction of these two master regulators of eye differentiation explains the morphological eye phenotype, but we noticed that embryonic transcription was already misregulated upstream of these factors. The pan-neural markers Nrp1 (Figure 2C) and N-CAM (Figure S5E), which are induced during gastrula stages, were also strongly reduced in double morphants. However, several key markers of embryonic patterning were not perturbed, such as the organizer genes Chordin, Goosecoid and Xnr-3 at gastrula stages (Figure S5B). The anteroposterior patterning of the central nervous system (CNS) appeared also to

be normal given the wild-type-like expression patterns of *Otx2* and *Krox20* in fore- and hindbrain territories, respectively (Figure S5C). These results provide first evidence that H4K20 di- and trimethylation serves to regulate distinct developmental genes in a selective manner.

Xenopus Suv4-20h activity is required for normal development

The specificity of the developmental phenotypes arising from xSuv4-20h enzyme depletion was validated by rescue experiments, in which we coinjected increasing doses of murine Suv4-20h1/h2 mRNAs together with the xSuv4-20h MO-mix. Due to sequence divergence, transcripts of the murine orthologs escape inhibition by the MOs targeting the frog mRNAs. Already 2 ng of murine Suv4-20h transcripts were sufficient to rescue the eye defect in two thirds of the double morphant embryos ($p < 0.0001$, Fisher's exact test). In most cases, the retinal neuroepithelium regained its circular structure and near normal size, as well as a central lens (Figure 2A). The rescue efficiency did not increase with higher concentrations of mouse transcripts (Figure 2B, columns 4–6). The number of melanophores was also increased at their proper sites under rescue conditions (Figure 2A). Furthermore, the expression domains of *Rx-1* and *Nrpl* (Figure 2C), as well as *Pax-6* and *N-CAM* (Figure S5D and S5E) were efficiently restored.

To test, whether this phenotypic rescue requires Suv4-20h proteins or their enzymatic activity, we generated catalytically inactive murine Suv4-20h protein variants (Figure S6A), based on structural predictions [25,26]. Unlike the wild-type proteins, neither variant restored the H4K20me3 mark at heterochromatic foci in Suv4-20h DKO MEFs (Figure S6B). When tested side by side with the wild-type enzymes, the mutants did neither increase the abundance of the H4K20me2 and -me3 marks in wild-type frog embryos (Figure S6C, compare lanes 1, 3 and 5), nor rescue H4K20 methylation levels in xSuv4-20h double morphants (Figure S6C compare lanes 1, 4 and 6), although being expressed at similar levels (Figure S6D). Consequently, the inactive variants also failed to rescue the eye and melanophore phenotype (Figure S7A–S7C and S7D, compare columns 2–4).

In the course of these experiments we noticed that overexpression of either frog or mouse Suv4-20h1 and h2 proteins never caused any obvious morphological or molecular changes in the embryos (Figures S8B, S8C and S9B, S9C), despite strongly enhanced H4K20me3 levels in bulk chromatin (Figures S8A and S9A). In particular, morphological landmarks such as eyes and melanophores formed normal in size, number and location under overexpression conditions. Expression domains of marker genes such as *Rx-1* and *Pax-6* were unaffected (Figures S8D and S9D). Thus, H4K20 di- and trimethylation is required for normal development, but excess deposition of these marks has no apparent phenotypic consequences.

XSuv4-20h enzymes are required for ectoderm formation

The apparent functional selectivity of the ubiquitously expressed enzymes encouraged us to test, whether xSuv4-20h HMTases control additional aspects of germ layer formation and patterning. Therefore, we compared the expression of key developmental regulatory genes in uni-laterally injected control-morphants *versus* xSuv4-20h double morphants by RNA *in situ* hybridisation (listed in Figure 3A).

Based on our previous results, we continued with genes involved in neurogenesis (Figure 3B). At the open plate stage, primary neurons are specified in three stripes next to the dorsal midline on each side. At this time, each stripe expresses the neural specific regulatory genes Neurogenin-related 1a (*Ngnr-1a*) and Delta-like

1, as well as the differentiation marker N-tubulin. The expression of these three genes was extinguished in almost all of the xSuv4-20h MO-injected sides, while being unaffected by control-MO (Figure 3B). In addition to these stripes, Delta-like 1 mRNA delineates the anterior border of the neural plate, and this domain was also extinguished (Figure 3B). In contrast, mesodermal expression of Delta-like 1 around the blastoporus remained unaffected in morphant condition (Figure 3B, arrow). Delta-like 1 and N-tubulin stripes were effectively rescued by coinjection of wild-type mSuv4-20h1/h2 mRNAs, while *Ngnr-1a* was restored in a broad, diffuse manner (Figure 3B, right column). Notably, inactive mouse Suv4-20h HMTases could not rescue N-tubulin expression (Figure S7E, middle column). At the same time, mesodermal control genes like *MyoD* were unaffected (Figure S7E, right column). Together, these results implicate xSuv4-20h enzymes in neuronal fate selection.

Next, we extended our analysis to marker genes expressed in other germ layers and territories (Figure 3C and Figure S5). The epidermal keratin gene *XK81* demarcates non-neural ectoderm and was expressed normally on the surface of morphant epidermis; however, due to a slight retardation in neural tube closure on the injected side, its expression appears asymmetric in anterior views. This may indicate an involvement of xSuv4-20h enzymes in morphogenetic processes during neurulation and/or neural crest specification. This phenotype led to a mild broadening of the neural plate markers *Sox2* (Figure 3C), *Sox3* and *Sox11* (Figure S5F) at apparently normal mRNA levels. Prior to these neural plate markers, a group of genes including *FoxD5*, *Geminin*, *Zic1*, *Zic2*, *Zic3* and members of the Iroquois family are induced in the prospective neuroectoderm and stabilize the neural fate by their regulatory interactions (reviewed in ref [27]). At midgastrula (*NF11*), *FoxD5* and *Geminin* did not respond to xSuv4-20h enzyme depletion (Figure S5F), but *Xiro1*, *Zic1* (Figure 3C), *Zic2* and *Zic3* (Figure S5F) mRNAs were strongly reduced. In contrast, key mesodermal factors such as *Xbra*, *MyoD* (Figure 3C) and *VegT* (Figure S5F), as well as regulators of endodermal differentiation like *Sox17* α and *Endodermin* (Figure 3C) were expressed normally in both morphants and in embryos overexpressing frog xSuv4-20h proteins (Figure S8E). Taken together these results demonstrate that xSuv4-20h HMTases are critical for neural development, but apparently dispensable for mesoderm and endoderm formation in *X. laevis*.

To further verify the specific role of *Xenopus* Suv4-20h enzymes in neural development, we considered two different approaches; in a first series of experiments we performed injections at 8-cell stage in the animal or vegetal pole blastomeres, selectively labelling cells predominantly belonging either to mesendoderm (vegetal injections, Figure 4A) or ectoderm (animal injections, Figure 4D). Vegetal pole blastomere injections led to no evident morphological and molecular phenotypes (Figure 4B and 4C). Conversely, animal injections reproduced the eye and melanophore phenotypes from half-injected embryos, while mesodermal and endodermal structures developed normally (Figure 4E). Consistent with the morphological defects, Delta-like 1 expression in the neural plate was suppressed, while *MyoD* and *Sox17* α genes were unaffected (Figure 4F). These results provide strong evidence that the neural and melanocyte phenotypes originate in the ectoderm.

As second approach we took advantage of animal cap (AC) explants, which form epidermis in isolation but can be neuralized by the BMP-inhibitor *Noggin*. Specifically, we tested whether the downregulation of xSuv4-20h HMTases prevented neural induction by *Noggin*. Without *Noggin*, wt and double morphant explants were positive for *XK81* and negative for *Nrpl*

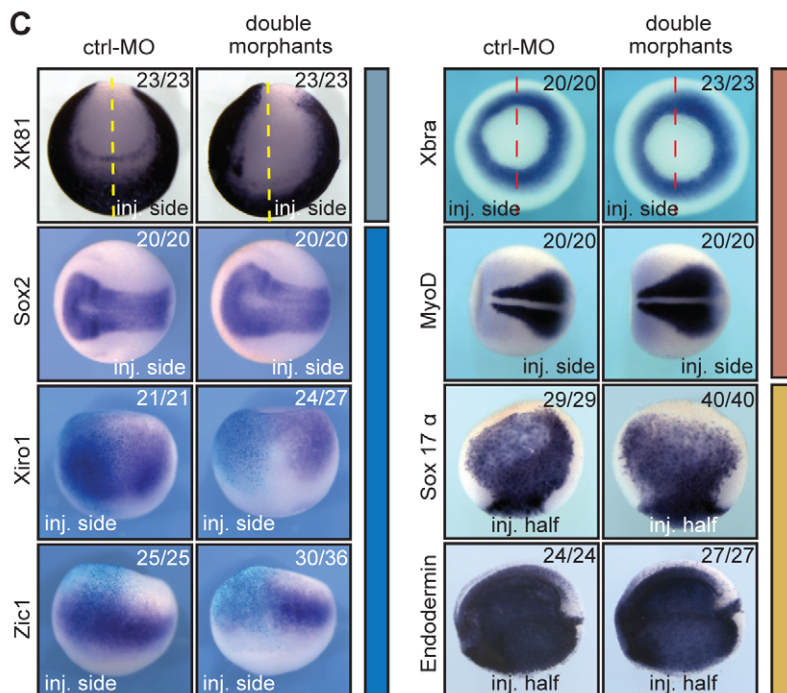
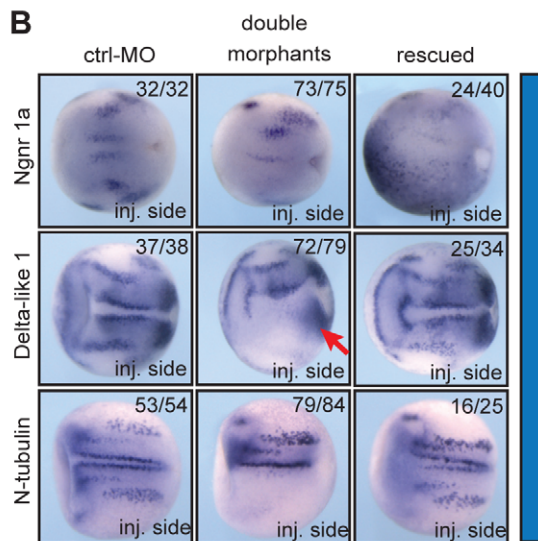
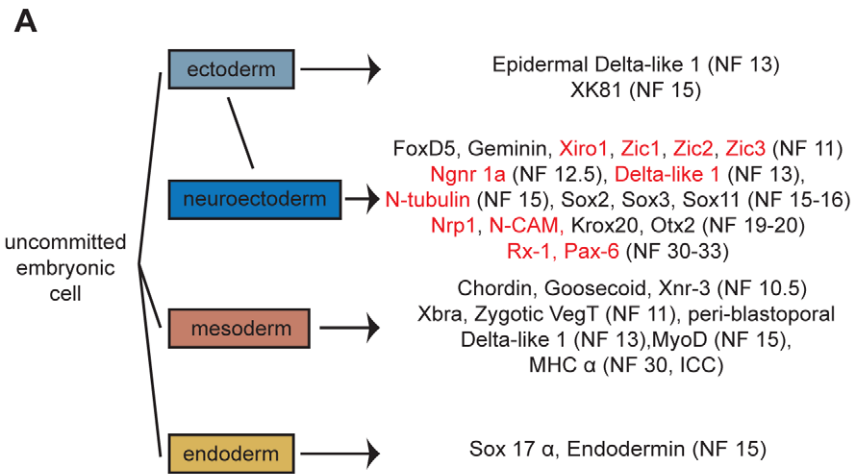


Figure 3. xSuv4-20h enzymes are required for differentiation of the neuroectoderm. (A) Schematic illustration of analysed markers of the different germ layers (germ layer colour code extended to *in situ* panels). Downregulated genes upon xSuv4-20h depletion are labelled in red. (B) Expression pattern of the neuroectodermal markers Ngnr 1a (NF12.5), Delta-like 1 (NF13), and N-tubulin (NF15). The pictures show dorsal views of the open neural plate with anterior to the left. (C) Expression patterns of XK81 (ectoderm), Sox2, Xiro1, Zic1 (neuroectoderm), Xbra, MyoD (mesoderm), Sox17 α , Endodermin (endoderm) in ctrl-MO injected or double morphant embryos. XK81 - anterior views with dorsal side to the top. Sox2 and MyoD - dorsal views, anterior to the left. Xiro1 and Zic1 - dorsal views; injected halves are lineage-traced by coinjection of LacZ mRNA and subsequent β -Gal staining (light blue). Xbra - vegetal view. Sox-17 α - internal stain from the injected side in bisected embryos, animal pole up. Endodermin - internal stain from the injected side in bisected embryos; anterior to the left.

doi:10.1371/journal.pgen.1003188.g003

(Figure 5A). They were also negative for Xbra, indicating absence of contaminating mesoderm. Noggin-mediated Nrp1 expression was clearly visible in wt caps, but strongly reduced upon coinjection of xSuv4-20h morpholinos, while XK81 expression was downregulated in both the samples (Figure 5A). Thus, double morphant caps are both refractory to neural induction and restrained in epidermal differentiation. However, they differentiate into mesoderm upon stimulation with Activin A just like control explants, as shown by immunostaining for muscle myosin heavy chain (Figure 5B). These results confirm the crucial role of xSuv4-20h enzymes in coordinating the formation of ectodermal tissues, and show that in the absence of the two enzymes neural induction is impaired.

XSuv4-20h enzymes are required for cell survival and proliferation

Loss of H4K20 di- and trimethylation is known to compromise DNA damage repair in mice and to partially block G1/S transition [24]. This prompted us to test, whether xSuv4-20h depletion affects apoptosis and cell proliferation in frog embryos. Immunostaining for activated Caspase3 revealed an increase in apoptotic cells on the injected side of double morphant embryos (Figure S10A). Coinjection of xBcl-2 mRNA, an anti-apoptotic factor, reduced the Caspase3 positive cells to levels of the uninjected control side, however, without re-establishing a proper Delta-like 1 and N-tubulin pattern in the double-morphant side. Overexpression of xBcl-2 mRNA alone had no effect on the expression of the tested markers (Figure S10A). Thus, although embryonic frog cells depleted for the H4K20me2/me3 marks become apoptotic at higher rate than wt cells, the absence of neurons in the double-morphant neural plate cannot be explained by selective cell death.

Double morphant embryos stained for the mitotic marker H3S10P, showed a two-fold reduction ($p = 0.0058$) in the number of proliferating cells at midneurula stage, compared to control morphant embryos (Figure S10B). This mild phenotype might be correlated with the observed increase in apoptosis. Since neural induction continues in frogs, even when cell proliferation is blocked from gastrulation onwards [28], it is unlikely that the nearly complete loss of N-tubulin positive neurons is brought about by this mild reduction in cell proliferation. Taken together, the main xSuv4-20h morphant phenotype represents not a selective loss of neuroblasts, but a block in neural differentiation.

XSuv4-20h double morphant frog embryos fail to silence Oct-25 transcription in sensorial ectoderm

So far, our analysis in xSuv4-20h morphant embryos has indicated a specific and selective loss of gene expression in ectodermally derived tissues. The earliest affected markers - Zic and Xiro genes - become induced at early gastrula stage and help establish the neural plate state [27]. At this time in frog development, embryonic cells in the animal hemisphere are still plastic and express members of the POU-V gene family - i.e. Oct-25, Oct-60 and Oct-91 - that encode paralogs of the mammalian

pluripotency regulator Oct4 [29,30]. Because Oct-25 and Oct-91 regulate germ layer differentiation in *Xenopus* [31–34], we investigated their expression (Figure 6A). Oct-25 is initially expressed throughout the animal hemisphere at early gastrula, but gets restricted to the presumptive floor plate (notoplate) by midneurula [31]. On the injected side of the vast majority of double morphants, however, Oct-25 expression was expanded from the notoplate down to the ventral midline. Interestingly, ectopic Oct-25 expression was restricted to the sensorial cell layer of the ectoderm, which contains neural and epidermal precursor cells, respectively (Figure 6A, sections). The Oct-60 gene, which is expressed during oogenesis, was not activated under these conditions. Oct-91 staining appeared normal in the majority of the embryos, although some showed a mild upregulation in double morphants as well (data not shown). The ectopic expression of Oct-25 is a specific consequence of xSuv4-20h enzyme depletion, because its normal pattern was re-established in morphants upon coinjection of mRNAs encoding wild-type, but not inactive, mouse Suv4-20h proteins (Figure S7E, left column). Notably, the selective derepression of the Oct-25 gene was also observed in double-morphant AC explants (Figure 6B), excluding indirect effects from non-ectodermal tissues.

We then performed qRT-PCR analysis to quantitate the relative changes in gene expression. It is frequently observed that embryo cohorts develop in slight asynchrony as a non-specific consequence of Morpholino injection, possibly obscuring transcriptional responses. To minimize this potential artifact, we analysed the RNAs of matching pairs of wt and xSuv4-20h depleted samples by dissecting embryos at early neurula stage (NF14) into uninjected and injected halves, based on the coinjected fluorescent lineage tracer (Figure S11A). As shown in Figure 6C, the Oct-25 mRNA is about three-fold higher in xSuv4-20h double-morphant halves ($p = 0.0123$), while being similar between control-morphant and uninjected halves. In the same sample, Oct-91 expression was unaffected (Figure 6C). We used this assay also to confirm the diminished expression of neural plate marker genes detected earlier by RNA *in situ* hybridisation. With the exception of Ngnr 1a, Nrp1 and N-tubulin, mRNA levels were clearly reduced in the morphant halves ($p = 0.0122$ and 0.0163 , respectively; Figure S11B).

To gain further information about the complexity of the underlying transcriptional misregulation, we performed transcriptome analysis in wild-type and double-morphant embryos, again dissecting embryos in corresponding pairs of injected and uninjected halves (Figure S12A). Six percent of the 11639 annotated probe sets present on the microarray were significantly altered in their expression as a consequence of xSuv4-20h enzyme depletion, about equally split into upregulated ($n = 319$) and downregulated ($n = 404$) probes (Figure S12B and S12C; for a complete list of the responding probesets see NCBI's GEO Series accession number GSE41256). This result suggests that the observed phenotypes in the double morphants originate from transcriptional misregulation of distinct genes, rather than from global, pleiotropic effects. Indeed, Oct-25 mRNA is also specifically upregulated in the microarray data set, where it is among the

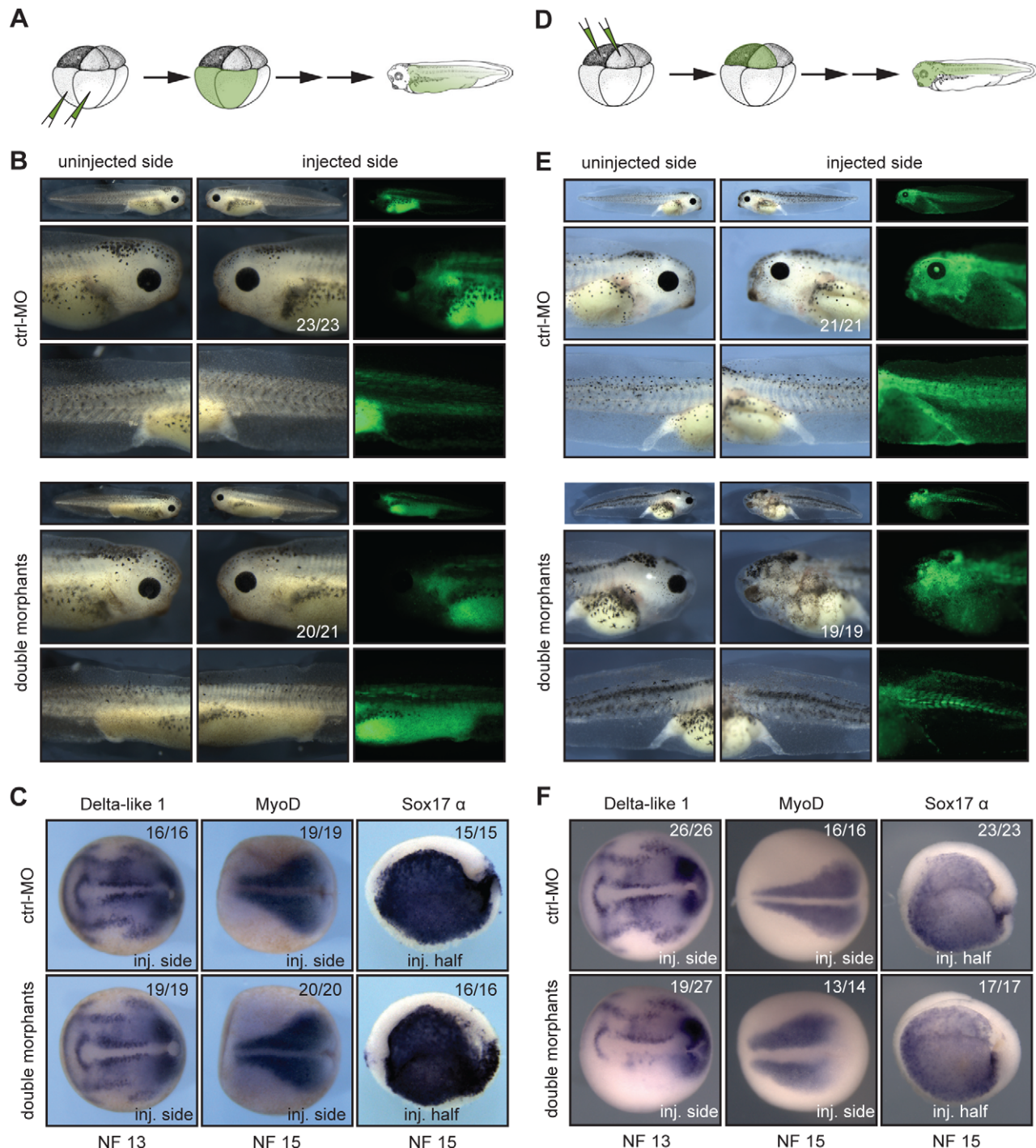


Figure 4. xSuv4-20h1/h2 enzymatic activity is required in the ectodermal germ layer. (A, D) Schematic illustrations of targeting microinjections into mesendodermal or ectodermal territories at 8-cell stage. (B) Injecting xSuv4-20h MOs into the mesendoderm causes no apparent morphological phenotype in the embryo. (C) Neural, mesodermal and endodermal marker genes are expressed normally. (E) xSuv4-20h MOs reduce eyes, cranial and trunk melanophores, when injected into the ectoderm. (F) Expression of all tested markers in mesoderm and endoderm is normal, except for Delta-like 1, whose expression specifically in the open neural plate is strongly reduced on the injected side. Global morphology was assessed at hatching stage (NF36), molecular markers at indicated stages during neurulation. Top row images in (B) and (E) depict whole embryos for overview.
doi:10.1371/journal.pgen.1003188.g004

ten most upregulated mRNAs in the double-morphant condition (Figure S12D).

The sustained expression of Oct-25 in xSuv4-20h morphant embryos fits the prediction of Oct-25 being a direct target of

H4K20me3 mediated transcriptional silencing. To test this assumption directly, we carried out chromatin-immunoprecipitation (ChIP) experiments with H4K20me3-specific antibodies at the neurula stage (NF15-16). For ChIP experiments we used *X.*

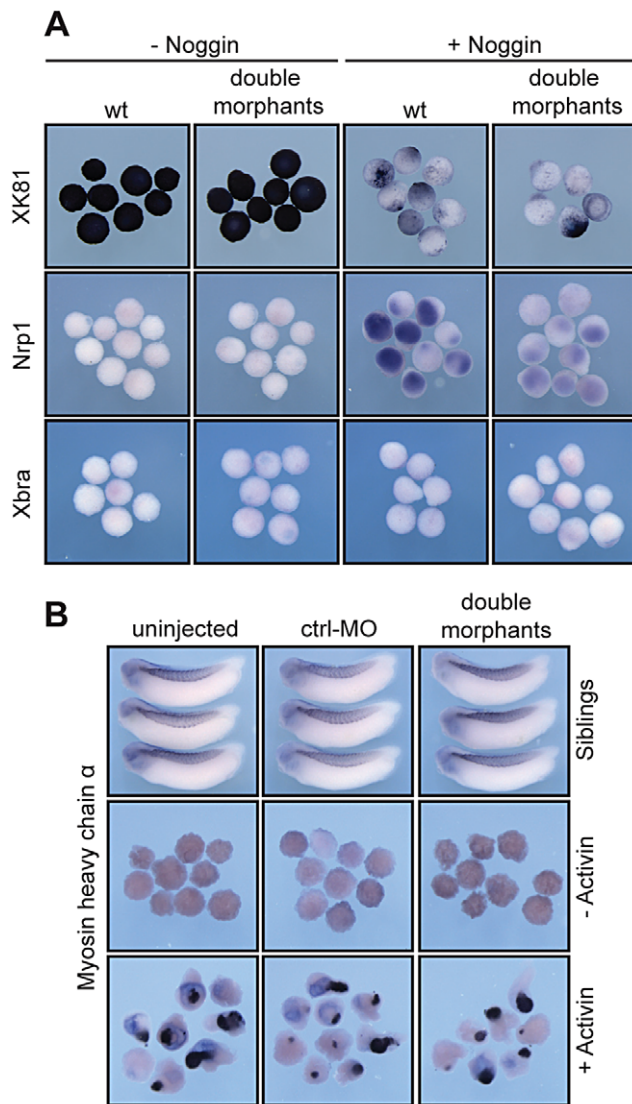


Figure 5. In vitro induction of xSuv4-20h double-morphant animal cap explants. (A) Noggin-dependent neuralisation. XK81, Nrp1 and Xbra expression is monitored in uninjected control caps and double-morphant caps with or without Noggin mRNA. Note that explants coinjected with xSuv4-20h MOs together with Noggin mRNA show reduced Nrp1 expression, but normal downregulation of XK81 mRNA. (B) Muscle induction by Activin A in uninjected, ctrl-MO injected, and xSuv4-20h double morphant animal caps. Top row demonstrates comparable expression of myosin heavy chain (MHC- α) in non-dissected sibling embryos.
doi:10.1371/journal.pgen.1003188.g005

tropicalis embryos, since the available genome sequence of this closely related frog species [35] allowed us to design primer amplicons for non-exon derived DNA sequences. RNA *in situ* hybridization performed on neurula stage *X. tropicalis* embryos, confirmed that the expression patterns of Oct-25 and N-tubulin were up- and down-regulated, respectively, to the same extent as observed for *X. laevis* (Figure S13). We retrieved the pericentromeric major satellite repeat sequence (MSAT3) as positive control amplicon for the experiment. Genic regions, which are H4K20me3-free and, thus, could be used as negative controls, are difficult to predict, since genome-wide analysis in mammalian cells reported only enrichment of this modification on pericentromeric and subtelomeric heterochromatin [36,37]. As negative

controls we considered: GAPDH, a constitutively expressed housekeeping gene; thyroid hormone receptor α (*thra*), a gene whose expression can be detected at neurula; and *thra*-induced *bzip* protein (*thibz*) that is expressed from metamorphosis on (Figure S14A). Statistical analysis of qRT/PCR data indicates that expression of GAPDH and *thra* was not significantly altered under the double-morphant condition (Figure S14B). Therefore, the relative H4K20me3 levels at these genes were defined as background, and compared to the levels on other loci (Figure S14A). The modification strongly decorated the pericentromeric MSAT3 repeat region (Figure 6D), as expected from the analysis in murine cells [21]. At the 5'UTR amplicon of the Oct-25 gene, H4K20me3 was significantly enriched compared to the control genes GAPDH ($p=0.0155$), *thra* ($p=0.0103$) and *thibz* ($p=0.0128$) (Figure S14A and Figure 6D). In a second set of experiments, we compared the abundance of H4K20me3 between wild-type and xSuv4-20h double-morphant embryos (Figure 6E). In morphants, the modification was selectively reduced at the 5'UTR amplicon of Oct-25 ($p=0.004$). Together, these ChIP experiments validate the 5' end of the Oct-25 gene as direct target of xSuv4-20h mediated transcriptional silencing.

Xenopus Oct-25 has been implicated in germ layer formation [32,34]. We wanted to know, whether the sustained expression of Oct-25 in xSuv4-20h morphants could cause the observed downregulation of early neural plate and neural differentiation markers. This question is difficult to address, since the role of Oct-25 in neural induction is ambiguous - both overexpression and morpholino knockdown inhibit neural differentiation [32,34]. Thus, Oct-25 acts in pleiotropic fashion, perhaps switching target genes or protein interaction partners. In a previous report [38], human Oct4 protein was shown by ChIP analysis to bind to promoters of early neural markers, including *Zic* and *Sox* genes. In order to link Xenopus Oct-25 mechanistically to these genes, we have mis-expressed constitutively activating and repressing Oct-25 fusion proteins in animal caps (Figure S15A). *Zic1*, *Zic3* and *Sox2* responded to the Oct-25 variants in a manner consistent with direct regulator/target gene interaction, i.e. they were hyperactivated by Oct-25-VP16 ($p=0.0143$; 0.0456; 0.01622, respectively) and suppressed by Oct-25-EnR ($p=0.0236$; 0.0167; 0.0231, respectively) compared to the uninjected sample. In line with this assumption, inspection of the *X. tropicalis* gene sequences detailed the presence of multiple Oct-25 DNA binding motifs within 2.0 Kb distance from the transcriptional start site for each of these genes (Figure S16). For the two *Zic* genes, which are misregulated in the forming neural plate of morphant embryos (Figure 3C and Figure S5F), we confirmed the misregulation by Oct-25 variants via RNA *in situ* hybridisation (Figure S15B).

Interestingly, *Sox2* expression was affected only in AC explants, but not in the double morphant embryos. This can be explained by considering two points: First, in animal caps levels and activities of the injected Oct-25 protein variants most likely exceed endogenous Oct-25 protein activity and regulate *Sox2* expression in a dominant fashion. Secondly, formation of neural tissue in the embryo requires inductive influences including FGF signalling [39], and *Sox2* transcription is stimulated by FGF8 [27], which is normally expressed in the mesoderm. Thus, the stimulating influence of FGF signalling on *Sox2* transcription in the embryo may neutralize the repressive influence from deregulated Oct-25 expression, while the repressive activity of the deregulated Oct-25 levels prevails in animal caps in the absence of FGF signalling.

The remaining genes either failed to respond to one of the two Oct-25 protein variants (*Zic2*, *Xiro1*), or did not respond (*Ngnr 1a*, N-tubulin). These observations suggest an indirect effect. While it is possible that additional factors that are misregulated in xSuv4-20h

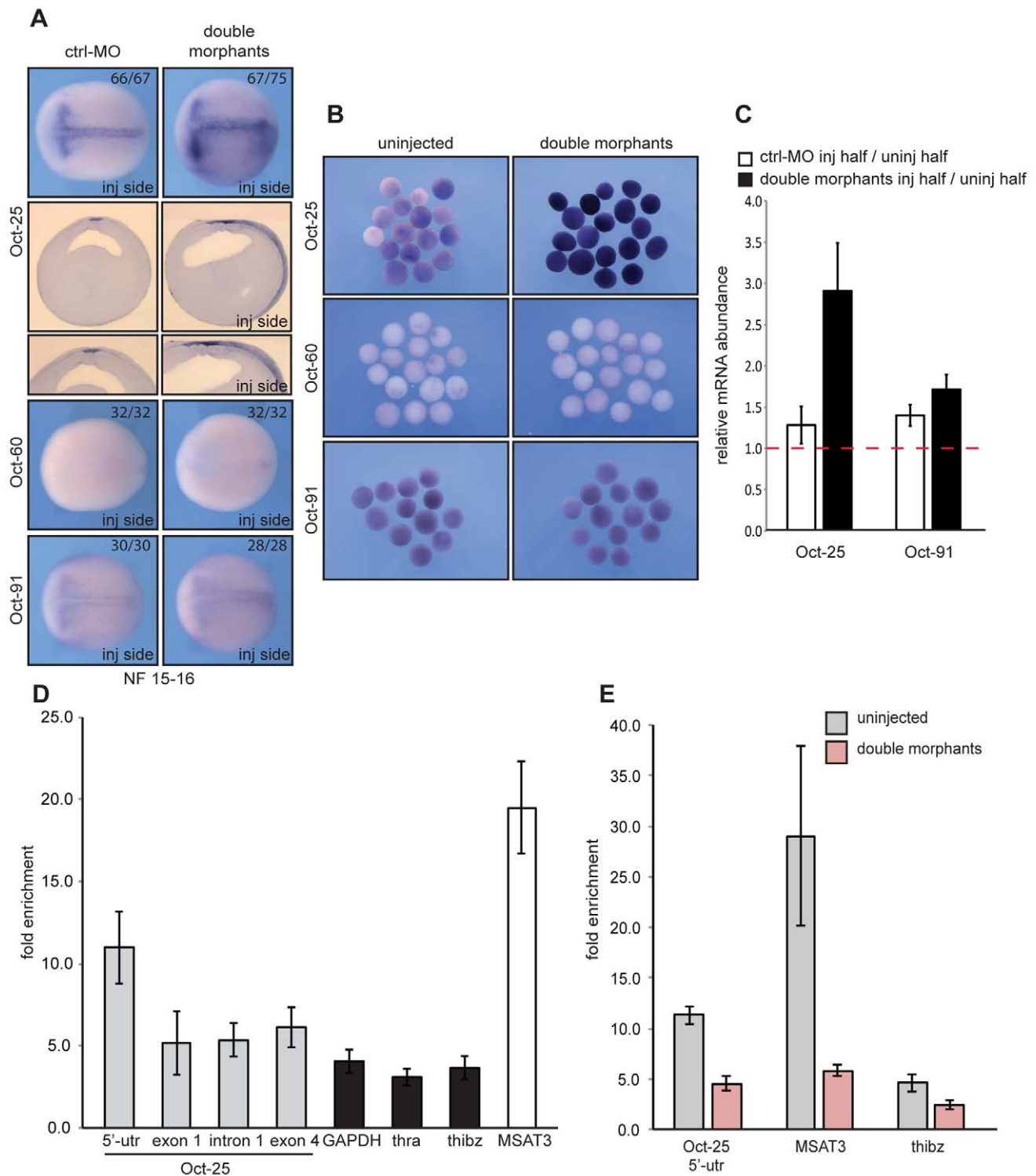


Figure 6. xSuv4-20h double morphants fail to silence Oct-25 expression in deep-layer ectoderm due to reduced H4K20me3 enrichment at Oct-25 promoter. RNA *in situ* hybridization analysis for Oct-25, Oct-60 and Oct-91 in embryos (A) and animal caps (B) for ctrl-morphants or xSuv4-20h double-morphant embryos, injected unilaterally at two-cell stage and fixed at midneurula stage (NF15). Injected sides were defined by coinjected Alexa-fluorescence prior to *in situ* hybridisation. (A) Dorsal views of stained embryos with anterior to the left. For Oct-25 Vibratome cross-sections, ctrl-MO and double-morphant embryos are shown. (B) Comparative expression analysis for Oct-25, Oct-60 and Oct-91 in animal caps from bilaterally injected embryos, fixed at midneurula stage (NF15). (C) qRT-PCR profiles for Oct-25 and Oct-91 in ctrl-MO and xSuv4-20h double-morphant embryos. Data represent normalized ratios of mRNA levels as means of four independent experiments, error bars indicate SEM. (D, E) Chromatin immunoprecipitation (ChIP) analysis on Oct-25, GAPDH, thra, thibz genes and major satellite repeat region 3 (MSAT3). (D) H4K20me3 levels on the indicated genes in normal embryos, and (E) in uninjected versus double-morphant embryos. Fold enrichment was calculated as the ratio between H4K20me3 precipitated material over negative control (No antibody sample). Data represent mean values of three to five independent experiments, error bars indicate SEM. doi:10.1371/journal.pgen.1003188.g006

morphants contribute to the neural phenotype, the combined results from ChIP experiments and Oct-25 variants define a pathway, in which xSuv4-20h enzyme dependent repression of Oct-25 is needed during gastrulation for proper neuroectoderm differentiation.

Deregulated Oct-25 expression in xSuv4-20h double morphants inhibits neural differentiation

To further analyse the mechanistic interaction between xSuv4-20h enzymes and Oct-25, we performed rescue experiments with triple-morphant embryos, in which synthesis of Oct-25 and xSuv4-20h proteins was simultaneously blocked (Figure 7). The Oct-25 morpholino that we used has been shown before to inhibit efficiently Oct-25 translation from both non-allelic gene copies [40]. Because global Oct-25 depletion inhibits the formation of anterior neural structures [40], we employed two different strategies for the triple-knockdown to circumvent this problem. In a first series of analysis we injected a single A1 blastomere of 32-cell stage embryos to target cells that predominantly contribute to the retina and brain. Also in this experimental series, the morphology of double morphant eyes was strongly affected (Figure 7A). 71% of the injected embryos showed a clear reduction of retinal pigment, the remainders often restricted to the dorsal-most portion of the eyecup. The majority of the eyes contained no lens (Figure 7C). When the downregulation of xSuv4-20h enzymes was coupled to a concomitant knockdown of Oct-25 (triple morphants), the percentage of embryos showing this defect was reduced to 49% ($p=0.0188$, Fisher's exact test). The retinal pigment was rescued in the triple morphants, whose eyes also regained a properly structured lens (Figure 7C). To confirm the morphological phenotypes, we investigated the basal neural gene expression in AC explants. The expression of a subset of genes involved in the establishment of the neural plate state (Zic1, Zic2, Xiro1, Sox2 and Sox3) was strongly reduced upon downregulation of xSuv4-20h enzymes at early neurula (NF14-15), compared to uninjected animal caps ($p=0.0068$; $p=0.0127$; $p=0.0113$; $p=0.0321$; $p=0.0037$, respectively). With the exception of Sox2, the simultaneous downregulation of xSuv4-20h enzymes and Oct-25, rescued neural gene expression. In fact, under the triple morphant condition most of these genes were expressed at higher levels than normal, suggesting that they are partly repressed by Oct-25 in unmanipulated explants (Figure 7D). Most importantly, the combined results of the two triple-knockdown experiments indicate that both morphological and molecular features of the xSuv4-20h double morphant phenotype can be rescued to a significant extent by reducing Oct-25 protein levels. This result firmly establishes that the sustained and elevated expression of Oct-25 protein is responsible for the neural differentiation defect of xSuv4-20h double-morphant embryos.

Murine Suv4-20h1/h2 double-knockout ES cells have elevated Oct4 levels in undifferentiating and differentiating conditions

Oct-25 plays multiple roles during early frog development, including interference with Activin/BMP-dependent mesoderm formation before gastrulation, and with neural induction during gastrulation [32,34]. A similar role is considered for its mammalian paralog Oct4, which is required for the pluripotent state of ES cells, but antagonizes ectodermal differentiation as soon as these cells exit pluripotency [30,41,42]. Although previous genome-wide studies of histone modifications in mammalian cells have not detected H4K20me3 on the Oct4 gene [36,37], this apparent similarity made us investigate Oct4 protein expression in wild-type and Suv4-20h1/h2 DKO murine ES cells. We tested

two independently derived DKO cell lines (B4-2 and B7-1), and compared them with two wild-type controls, i.e. wt26, an isogenic ES cell line, and the well-characterized GSES-1 cell line [43]. All four cell lines formed comparable ES cell colonies in LIF-containing medium (Figure 8A and Figure S17B), although the two DKO lines grew slightly slower. Upon aggregation they formed embryoid bodies, which were clearly smaller than those of the wild-type lines, both at day 2 and day 6 of differentiation (Figure 8A and Figure S17). After replating the differentiated cells for one day, the two DKO lines frequently formed again colonies resembling undifferentiated ES-cells (day 7 in Figure 8A and Figure S17B). To obtain a quantitative measure of Oct4 gene expression, we fixed and stained the four cell lines before (day 0) and during (day 6) differentiation for Oct4 protein and subjected equal cell numbers to FACS-analysis. The Oct4 signals were quite similar between wt26 and GSES-1 cells, as they were between the two DKO lines. In contrast to the wild-type cell lines, however, the signals of the DKO lines were reproducibly shifted to the right (Figure 8B and Figure S17C). Based on normalized median fluorescence intensity, the two DKO lines contained approximately three-fold higher Oct4 protein amounts than the wild-type lines at day 0 ($p=0.00604$), and still two-fold more at day 6 ($p=0.01266$) ($n=3$; Figure 8C and Figure S17B). We conclude that Oct4 expression is being reduced during differentiation in Suv4-20h1/h2 DKO cells. However, these cells have higher Oct4 levels in the undifferentiated state, and maintain higher levels during differentiation in comparison to wild-type cells.

Oct4 protein levels are known to be tightly regulated [1] and to influence lineage decisions during ES cell differentiation [41,42]. We therefore investigated the biological significance of the elevated Oct4 protein levels in Suv4-20h DKO ES cell lines. Unfortunately, the applied EB differentiation protocol promotes predominantly mesendodermal differentiation, which prevented the analysis of neural markers. Nevertheless, we performed FACS analysis on wt and Suv4-20h DKO cell lines stained for the chemokine receptor 4 (CXCR4) protein, whose expression indicates mesendoderm induction in embryoid bodies. At day 6 of differentiation, wt cell lines showed a robust increase in CXCR4 positive cells compared to day 0 (Figure 8D and data not shown). In contrast, both Suv4-20h DKO cell lines contained a significantly lower percentage of CXCR4 positive cells at day 6 when compared to the wild type cell lines ($p=0.03255$; Figure 8). We also noted that replated wt cells frequently formed autonomously beating areas at differentiation day 14 (see Video S1), indicating functional cardiomyocyte formation, while contracting areas were never observed in the Suv4-20h DKO cells (Video S2; $n=4$ experiments). Finally, qRT-PCR analysis indicated a reproducible and statistically significant shift in mesendoderm gene expression in the DKO ES cells, which show enhanced induction of FoxA2 ($p=0.00706$) and reduced levels of Gata4 ($p=0.00037$), compared to the wt ES cell lines (Figure S17D). Together, these results reveal a compromised and biased differentiation capacity for Suv4-20h DKO ES cell lines, and provide an entrypoint for further experimentation in the murine system.

Discussion

In this study, we have investigated the developmental functions of the histone- methyltransferases Suv4-20h1 and h2 during frog embryogenesis, which are responsible for the establishment of the H4K20 di- and trimethylated states. These modifications have been implicated in heterochromatin formation, DNA damage repair and G1/S-transition [21,24] and are also involved in

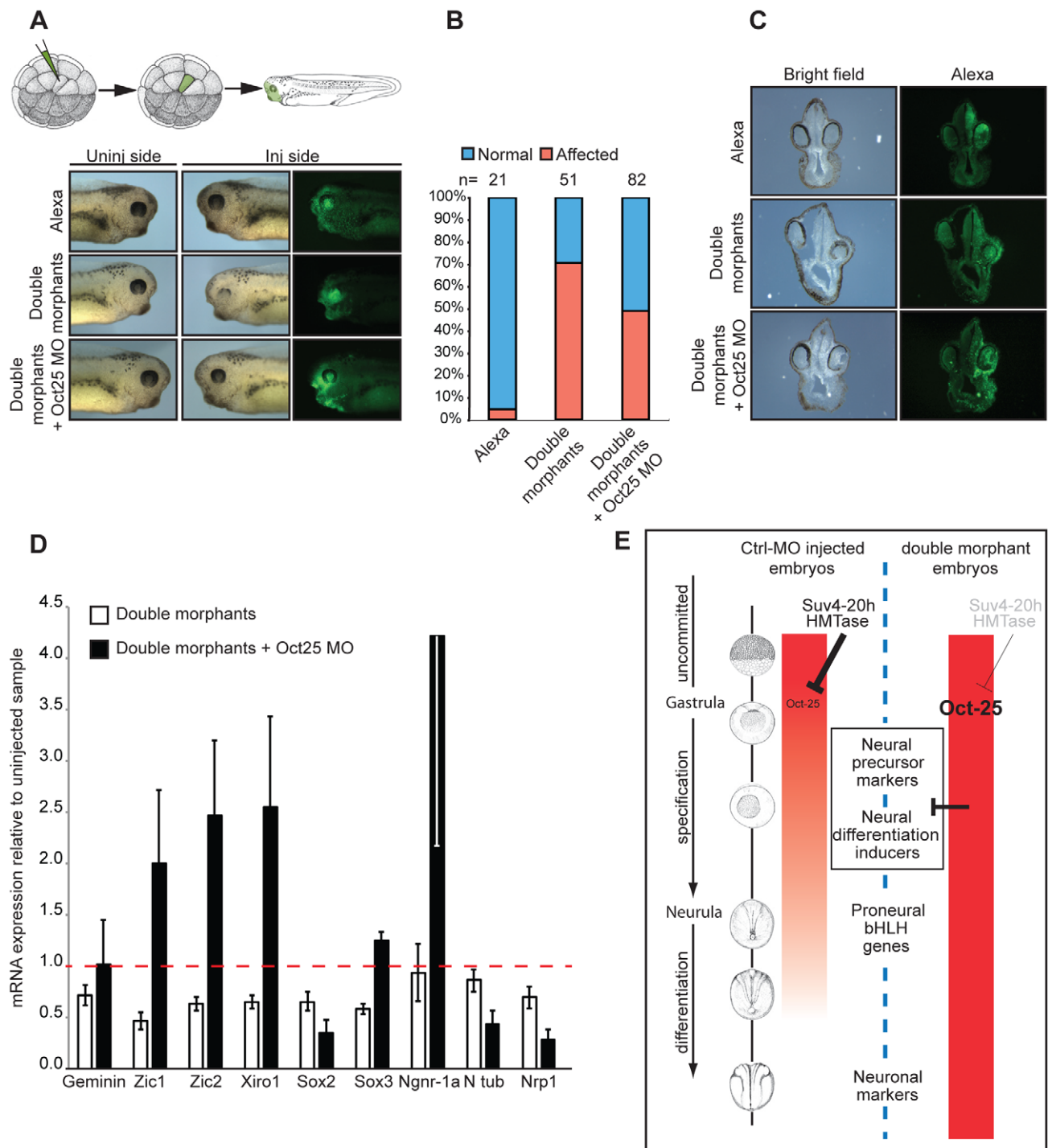


Figure 7. Oct-25 knockdown in double-morphant embryos rescues xSuv4-20h phenotypes. (A) Schematic illustration of targeting microinjections of tadpoles injected into the A1 blastomere at 32-cell stage, and morphological phenotypes of representative embryos (NF35-37) from cohorts injected with Alexa, xSuv4-20h MOs (double morphants) and double morphants plus Oct-25 MO. (B) Penetrance of the eye phenotype. Data from three independent experiments; n = total number of embryos scored. (C) Vibratome cross-sections of representative embryos injected as in panel (A). (D) qRT-PCR profiles for the indicated genes in double morphants and double morphants plus Oct-25 MO animal cap explants at NF 14–15. Data represent normalized mRNA levels as mean of three to four independent experiments; error bars indicate SEM. (E) Model for *Xenopus* Suv4-20h1/h2 enzyme function during neuroectoderm differentiation. A global increase in H4K20me3 reduces widespread Oct-25 expression in the animal hemisphere during gastrulation as a prerequisite for neural induction. In H4K20me3 depleted morphant embryos, Oct-25 expression persists in the ectodermal stem cell compartment (sensorial cell layer), interfering with the transcriptional activation or activities of key regulators of the neural plate state and neurogenesis. Additional genes that are deregulated like Oct-25 in xSuv4-20h morphant embryos, may also contribute to impaired ectoderm differentiation.

doi:10.1371/journal.pgen.1003188.g007

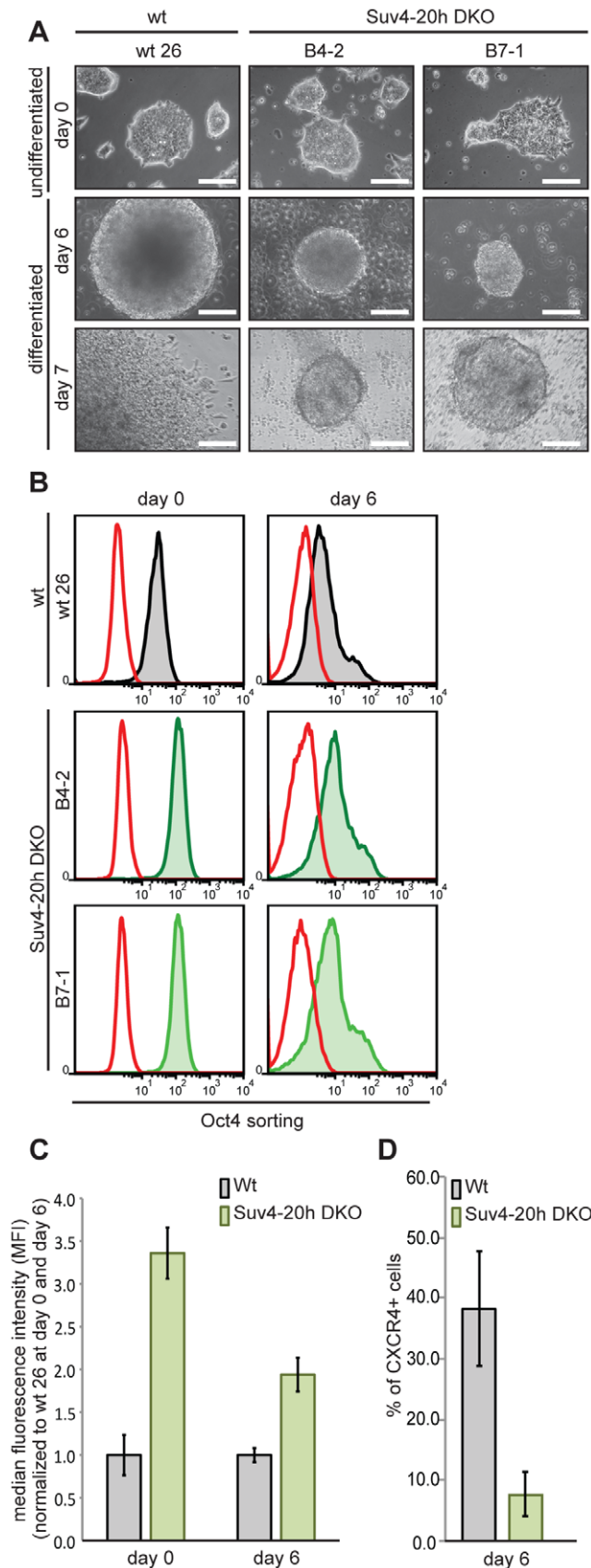


Figure 8. Suv4-20h double-null ES cells have elevated Oct4 and lower CXCR4 protein levels before and during differentiation. Wild-type and Suv4-20h DKO ES cells were grown undifferentiated in

LIF-containing medium, or differentiated *in vitro* by embryoid body formation. Wt26 - isogenic wild-type ES cell line; B4-2, B7-1 - independent Suv4-20h DKO ES cell lines. (A) Morphological appearance (scale bar 100 μ m). Top row: undifferentiated ES cells (day 0); middle: embryoid bodies at day 6; bottom: cells from embryoid bodies, replated for 24h. (B) Before (day 0) and during (day 6) differentiation, cell lines were stained for Oct4 protein and 2×10^4 cells from each condition were subjected to FACS analysis. Red graph: fluorescence of non-specific isotype control; black and green graphs represent the Oct4 protein levels in wild-type and Suv4-20h DKO ES cell lines, respectively. (C) Suv4-20h DKO cells have higher Oct4 protein levels compared to wild-type ES cells and maintain these during differentiation. Median fluorescence intensity was calculated from data in panel (B), error bars indicate SEM. (D) Suv4-20h DKO cells show a reduction in the percentage of CXCR4+ cells at differentiation day 6. Data represent normalized values of percentage of CXCR4+ cells as means of three independent experiments, error bars indicate SEM. doi:10.1371/journal.pgen.1003188.g008

transcriptional regulation [44,45]. Our experiments identify a specific and selective role of xSuv4-20h HMTases in the formation of the ectodermal germ layer through control of mRNA expression of key regulators of the neural plate state and neuronal differentiation circuits. Indeed, our results indicate for the first time that H4K20me3 controls transcription in a rather gene-specific manner. The mRNA profile of double morphant embryos shows appr. 6% of the annotated probesets to be misregulated, when H4K20me3 levels have been reduced to appr. 25%. About half of the responding mRNAs are transcriptionally upregulated and, thus, their genes may qualify as being directly controlled by H4K20me3 deposition. Surprisingly, our molecular analysis revealed that xSuv4-20h enzymes are required to restrict the expression of the pluripotency-associated Oct-25 gene during gastrula and neurula stages. In the absence of proper H4K20me3 deposition, the Oct-25 gene becomes transcriptionally derepressed and interferes with neural differentiation. The successful rescue in double-morphant embryos by the simultaneous inhibition of Oct-25 translation establishes this pathway formally. At least in *Xenopus*, the regulatory interaction between xSuv4-20h enzymes and Oct-25 is needed for embryonic cells to exit the pluripotent state and differentiate as neuroectoderm.

The genetic interaction between Suv4-20h enzymes and POU-V genes appears also to be conserved in mouse ES cells, although the H4K20me3 mark has not yet been detected on the Oct4 gene locus. To this point, we have shown that Suv4-20h DKO ES cells contain significantly elevated Oct4 protein levels, compared to wt ES cells. During ES cell differentiation the mammalian Oct4 gene is known to become repressed by a battery of epigenetic mechanisms including DNA methylation, incorporation of somatic linker histones and repressive histone modifications (H3K9me3/H3K27me3), which cooperate to achieve chromatin compaction of the Oct4 gene locus [46]. Our finding that Oct4 protein levels are increased in the DKO ES cells both before and during differentiation actually suggests that Suv4-20h enzymes regulate mammalian Oct4 transcription in a way that is at least partly independent from the other repressive mechanisms targeting this locus.

Our results in *Xenopus* rest predominantly on loss of function analysis, achieved by morpholino-mediated knockdown of endogenous xSuv4-20h protein translation. Specifically, we have shown that our antisense oligonucleotides block translation of xSuv4-20h1 and h2 isoforms *in vitro*, and significantly decrease H4K20me2 and -me3 levels *in vivo*, without altering the bulk

abundance of other repressive histone marks such as H3K9me3 and H3K27me3. The morpholinos produced specific phenotypes, which were rescued on the morphological and molecular level by RNA-born co-expression of heterologous xSuv4-20h enzymes and, thus, originate from deficient H4K20me2/me3 states.

While xSuv4-20h double morphant embryos showed consistent phenotypes at high penetrance, we were surprised to see that H4K20me2 and -me3 states could be quantitatively increased in frog embryos without any obvious morphological or molecular consequences (Figures S8 and S9). This result can be explained considering first of all the higher stability of the knockdown by non-degradable morpholinos compared to the transient protein upregulation by RNA injection; secondly, demethylation of higher-methylated states may occur rather rapidly through H4K20me2 and me3 demethylases at specific sites, where H4K20me1 is required, e.g. Wnt/ β -Catenin inducible genes [47]. However, we did not observe evidence for compromised transcription of Wnt target genes under overexpression (Figures S8 and S9) or morphant condition (Figure S5). Since mono- and dimethylated H4K20 states are quite abundant modifications in *Xenopus* embryos (30–40% each; see ref. [13]), it is most likely the loss of H4K20 trimethylation, which interferes with normal development.

XSuv4-20h double-morphant embryos were frequently defective for eye and melanocyte differentiation, indicating a prominent impairment of neuroectodermal differentiation. This selectivity is surprising, given that the two HMTases are expressed throughout the entire embryo (Figure S2). As a matter of fact, the phenotypes originate in the neuroectoderm, as shown by targeted injection into animal or vegetal blastomeres of 8-cell stage embryos (Figure 4). A large panel of marker genes that were investigated by RNA *in situ* hybridisation indicates that mesodermal and endodermal gene expression patterns are not perturbed by xSuv4-20h enzyme depletion (Figure 3A). This includes markers, which are required for specification of embryonic axes and formation and patterning of the mesendodermal germ layers (Figure S5). We also note that morphant animal cap explants were refractory to Noggin-dependent neural induction, but could be induced to differentiated skeletal muscle by a mesoderm inducing signal (Figure 5). We therefore assume that a major function of xSuv4-20h enzymes lies in the transcriptional control of genes that coordinate and execute neuroectodermal differentiation. Consistent with this hypothesis, many of the genes that we found downregulated in xSuv4-20h morphants, are key regulators of eye development (Rx-1, Pax-6), neuronal differentiation (Ngnr 1a, Delta-like 1) or regulators of neural competence and neural plate state (Zic-1, -2, -3, Xiro-1, Nrpl, N-CAM; [27]).

While these molecular results explain the overt morphological phenotypes in a consistent manner, it should be noted that these HMTases are clearly involved in additional cellular aspects. The mild reduction in mitotic cells and the increased apoptotic rate of morphant embryos (Figure S10) is reminiscent of findings in Suv4-20h1/h2 DKO MEFs, which are less resistant to DNA damage and compromised at the G1/S checkpoint [24]. The data reported here indicates a need for deeper analysis of the regulatory impact of Suv4-20h enzymes on transcription in both mammals and non-mammalian vertebrates.

According to current models, xSuv4-20h enzymes mediate transcriptional repression, based on the enrichment of the H4K20me3 mark on heterochromatic foci. Genes that are regulated by these enzymes should therefore become derepressed under loss of function condition. Following this logic, many of the genes, which are misregulated in morphant frog embryos, would be classified as indirect targets, since they were downregulated.

One very notable exception, which we have validated as direct target, is Oct-25 (Figure 6). Oct-25 is induced broadly in the animal hemisphere at the blastula/gastrula transition, before it becomes restricted to the notoplate at neurula stages [31]. Oct-25 plays multiple roles during early frog development, including interference with Activin/BMP-dependent mesendoderm formation before gastrulation, and with neural induction during gastrulation [31,32,34]. Our study reveals now a new function for Oct-25, namely to control the transit from a pluripotent cell to a neural cell that differentiates, when Oct-25 expression has faded. As depicted in our model (Figure 7E), this function depends on the precise dose and duration of Oct-25 transcription, which is controlled by the level of H4K20me3 deposition on the first exon of the Oct-25 gene through xSuv4-20h enzymes. As we have shown here, deregulated transcription of Oct-25 in double-morphant embryos elicits massive consequences on the differentiation of neuroectodermal organs and cell types. We have traced back the origin of the malformations to the gastrula stage, when a gene network, defining the neural state, become perturbed by Oct-25. Some members of this network are good candidates for direct regulation through Oct-25 (e.g. Zic and Sox genes). However, since Oct-25 transcription persists ectopically at least until the mid-neural fold stage in the ectoderm, subsequent gene cascades involved in regional differentiation of the neuroectoderm could also be directly misregulated by Oct-25.

The specific and selective deregulation of Oct-25 transcription in a precise spatial domain, i.e. the sensorial cell layer of the ectoderm, implies a very intriguing role for xSuv4-20h enzymes. This domain contains not only the uncommitted precursors of neuronal and epidermal cell types, but – with regard to the involuting marginal zone – includes also mesodermal and endodermal precursor cells. The observed derepression of Oct-25 in this domain may thus reflect a conserved mechanism, by which Suv4-20h enzymes control pluripotency in the embryo. As discussed above, we have found Oct4 protein to be increased in two independent Suv4-20h double knockout ES cell lines under LIF-maintained self-renewal conditions, when compared to wt ES cells (Figure 8 and Figure S17). The DKO cell lines also maintain higher Oct4 levels during differentiation than wt ES cells, although their Oct4 levels get diminished in the course of 6 days. Recent data from several labs suggest that the pluripotency regulators Sox2 and Oct4 guide ES cells towards specific germ layer differentiation programs, when they exit the pluripotent state [41,42]. Indeed, our findings are in agreement with Thomson and colleagues, who describe Oct4 to antagonize ectodermal specification and to direct mesendodermal cell fate decisions. The conserved Suv4-20h-dependent restriction of Oct4 expression may thus contribute to the germ-layer specification of embryonic cells, when they exit the pluripotent state.

Materials and Methods

Ethics statement

Animal work has been conducted in accordance with Deutsches Tierschutzgesetz; experimental use of *Xenopus* embryos has been licensed by the Government of Oberbayern (Projekt/AK ROB: 55.2.1.54-2532.6-3-11).

Expression constructs and *in vitro* transcription

Full length *X. laevis* Suv4-20h1a (NM_001092308) and Suv4-20h2a (NM_001097050) cDNAs in pCMV-SPORT6 were provided by ImaGenes. Capped mRNAs were synthesized *in vitro* with SP6 RNA-Polymerase after HpaI linearization. Both cDNAs were subcloned via XhoI/EcoRI sites into pBluescript II KS to

generate digoxigenin-labelled antisense probes with T3 RNA-Polymerase. *Xenopus* Bcl-2, Oct-25-VP16 and -EnR constructs were transcribed with SP6 RNA-Polymerase from NotI- (Bcl-2 and Oct-25-VP16) and SacII- (Oct-25-EnR) linearized pCS2+ plasmids, respectively. Mouse Suv4-20h1 and h2 enzymes were transcribed with SP6 from PvuI-linearized pCMVmyc-constrcuts [24]. Enzymatically inactive mouse Suv4-20h HMTases were generated via PCR-mutagenesis (see Text S1, Table S1 for primers). Synthetic mRNAs were injected in the animal pole of two-cell stage embryos at 2, 3 or 4 ng per embryo. Rescue experiments with wt and mutated mRNAs were performed with 3 ng of a 1:1 mix of wt or mutated Suv4-20h1 and h2 mRNAs, injected into the animal pole of a single blastomere at two-cell stage. *Xenopus* Bcl-2 mRNA was injected unilaterally in the animal pole of two-cell stage embryos at 800 pg per embryo. *Xenopus* Oct-25-VP16, -EnR mRNAs were injected in the animal pole of two-cell stage embryos at 100 pg per embryo.

Cell culture, microscopy, and FACS analysis

Mouse embryonic fibroblasts (MEF) wild type and Suv4-20h DKO cells [24] were cultivated in High Glucose DMEM with L-Glutamine and sodium pyruvate, complemented with 10% FCS, β -mercaptoethanol, non essential amino acids and penicillin/streptomycin in a 37°C incubator at 5% CO₂. Lipofectamine 2000 (Invitrogen) was used for the transfection of plasmid DNAs. Immunofluorescence analysis was performed as described in the Text S1.

Mouse ES cells were cultivated on gelatinized plates in High Glucose DMEM with L-Glutamine and sodium pyruvate, complemented with 15% FCS, 0.1 mM β -mercaptoethanol, non essential amino acids, penicillin/streptomycin and LIF. Cells were maintained at 37°C in a humidified atmosphere of 5% CO₂. ES *in vitro* differentiation and FACS analysis were carried out as described [43]. The incubation steps with the primary Oct4 (1:250, Abcam) or CxCR4 (1:50, BD Pharmingen) antibody and subsequently a FITC-conjugated secondary antibody (1:250, Invitrogen) were performed at RT for 45 min with two washing steps after each antibody incubation. For the isotype controls purified, IgG was used instead of the Oct4-antibody. All FACS analyses were performed with an Epics XL (Beckman-Coulter) using the analysis software FlowJo.

Morpholino oligonucleotides

Translation-blocking Morpholino oligonucleotides targeting *Xenopus* Suv4-20h1 (*X.laevis* and *X.tropicalis*: 5'-GGATTTCGCC-CAACCACTTCATGCCA-3'), *Xenopus* Suv4-20h2 (*X.laevis*: 5'-TTGCCGTCAACCGATTGAACCCAT-3'; *X.tropicalis*: 5'-CCGTCAAGCGATTGAACCCATAGT-3') and *Xenopus* Oct-25 (*X.laevis*: 5'-TTGGGAAGGGCTGTTGGCTGTACAT-3') mRNAs were supplied by Gene Tools LLC. Each Morpholino recognizes the two non-allelic isoforms of each gene in *X.laevis* (see Figure S3A, S3B). GeneTools' standard control Morpholino was used to monitor non-specific effects. Morpholino activity was tested by *in vitro* translation (SP6-TNT Kit, Promega), adding 2 pg of control Morpholino or 1 pg of Suv4-20h1 and/or h2 Morpholinos per TNT reaction. Unless stated otherwise, embryos were injected at a dose of 60–80 ng per embryo (30–40 ng each of Suv4-20h1 and h2 Morpholinos, or 60–80 ng control Morpholino per embryo). For 8-cell stage experiments, morpholinos were injected in two neighbouring, animal or vegetal blastomeres on one side of the embryos, at half the dose (i.e. 40 ng total). For morphological epistasis experiments, *Xenopus* Suv4-20h1 and h2 Morpholinos (5 ng each per embryo) and Oct-25 Morpholino (1 ng per embryo) were injected into A1 blastomere at 32-cell stage.

Embryo handling

Xenopus laevis eggs were collected, fertilized *in vitro*, and handled following standard procedures; embryos were staged according to Nieuwkoop and Faber (1967). The embryos were injected with maximally 10 nl volume. When required, they were sorted into left side or right side injected cohorts before fixation, based on the coinjected lineage tracer Alexa Fluor-488 Dextran (Invitrogen). Alkaline-phosphatase stained and refixed embryos were either sectioned after embedding in paraffin (10 μ m), or in gelatine/albumin mixture supplemented with 25% glutaraldehyde before sectioning (30–50 μ m) with a Vibratome 1000 (Technical Products International, INC.) as described [48]. Animal caps were manually dissected at NF9 and transferred singly into wells of a 96-well plate, coated with 1% agarose and filled with 1X Steinberg's solution, 0.1% BSA with or without Activin A (1:10 diluted conditioned cell culture supernatant). For neural induction, embryos were injected into the animal pole with Noggin mRNA (60 pg per embryo) alone or together with xSuv4-20h1 and h2 morpholinos (40 ng each per embryo) at two- to four-cell stage. For mesoderm induction, embryos were injected animally 4 times with 2.5 nl of control morpholino (80 ng per embryo) or a mix of xSuv4-20h1 and h2 morpholinos (40 ng each) at two or four cell stage. For Oct-25-VP16 and -EnR overexpression experiments, embryos were injected animally 4 times with 2.5 nl of each mRNAs (100 pg per embryo). For epistasis experiments on animal caps, embryos were injected 4 times with 2.5 nl of xSuv4-20h1 and h2 Morpholinos (40 ng each per embryo) and Oct-25 Morpholino (30 ng per embryo) at two or four cell stage.

Analysis of histone modifications in *Xenopus* embryos

Nuclei extraction from *Xenopus* embryos and mass spectrometry analysis of histone modifications were performed as described [13]. Histone marks were quantitated as relative abundances of a specific modification state as a fraction of the amount of all modifications found for this peptide (for details see ref 13).

RNA *in situ* hybridization and immunocytochemistry

Whole-mount RNA *in situ* hybridizations were performed as described (Sive et al. 2000). Embryos were photographed under bright light with a Leica M205FA stereomicroscope. The following antibodies were used for immunocytochemistry: H3S10P antibody (1:300, Upstate Biotechnology), active Caspase3 antibody (1:500, Promega), and myosin heavy chain antibody MF20 (1:100 hybridoma cell culture supernatant), anti-mouse or anti-rabbit alkaline phosphatase-conjugated secondary antibodies (1:1000, Chemicon).

Western blots and immunostaining

Embryonic histones were purified via acidic extraction of nuclei as described [13], size-separated by SDS-PAGE and blotted onto PVDF membranes (Roth). Membranes were blocked with 3% BSA (Roth) in PBS and subsequently incubated o/n at 4°C with polyclonal rabbit antibodies against H4K20me1 (1:6000), H4K20me2 (1:1000), H4K20me3 (1:500) [21,24] and pan H3 (1:25000, Abcam). Infrared (IR) 680 or 800 conjugated Goat anti Rabbit IgG (1:5000, Li-Cor) were used as secondary antibodies (incubation o/n at 4°C). Signals were detected with an ODYSSEY Infrared Imaging System. To extract exogenous myc-tagged fusion proteins embryos were treated as described in the Text S1. Proteins were separated by SDS-PAGE, BSA-blocked PVDF membranes were incubated o/n at 4°C with anti-myc 9E10 antibody (1:50), followed by anti-mouse HRP- conjugated antibody (1:3000, Jackson ImmunoResearch). Proteins were

detected with ECL plus western blotting detection reagents (GE Healthcare). Histological sections were stained with pan H3 (1:2000, Abcam), H4K20me1 (1:5000), H4K20me2 (1:2000), H4K20me3 (1:5000) antibodies [24].

Quantitative RNA analysis

Total cellular RNA was isolated with TRizol (Qiagen) and phenol/chloroform extraction. On-column RNA clean-up, including a DNase digestion step, was performed using RNeasy-Mini-Kit (Qiagen). Samples for qRT-PCR and microarray profiling were collected as described in the Text S1.

Microarray expression analysis

Microarray data were processed using R/Bioconductor (www.bioconductor.org). If not indicated otherwise, we used standard parameters in all functions calls. Expression values were calculated using 'gcrma'. Probe sets were kept for differential expression analysis if there were more 'present' calls (calculated using 'mas5calls') in one of the treatment groups than non-'present' calls, if their expression level variance was higher than 0 across all arrays and if the probe set had an Entrez identifier annotation according to the Entrez database with a date stamp of 2011-Mar16. One gene to many probe set relationships were resolved by retaining only the probe set with the highest variance across all arrays. Differential expression statistics were obtained using a linear model (library 'limma'). A significant response was defined if the local false discovery ('locfdr' package) rate calculated on the moderated t statistic was smaller than 0.2. The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE41256 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE41256>).

ChIP experiments

ChIP experiments were performed using *Xenopus tropicalis* as described [49], with minor changes (see Text S1 for details).

Identification of Oct-25 binding sites

A published weight matrix (PMID:17567999) was used to scan 2 kb upstream regions of selected *X. tropicalis* genes (Xenbase version 7.1) for binding site occurrence. Scanning was performed using RSA matrix-scan (PMID:18802439) with default parameters.

Supporting Information

Figure S1 *Xenopus laevis* versus *Mus musculus* Suv4-20h proteins sequence alignment. Amino acid sequence alignment for *Mus musculus* (Refseq. NM_001167885.1) versus *Xenopus laevis* Suv4-20h1 (Refseq. NM_001092308) (A) and *Mus musculus* (Refseq. NM_146177.2) [24] versus *Xenopus laevis* Suv4-20h2 (Refseq. NM_001097050) proteins (B). (C) Amino acid sequence conservation (% identity) of the SET domain between mouse and *Xenopus* proteins. (PDF)

Figure S2 *Xenopus laevis* Suv4-20h expression during early development. XSuv4-20h1 (A) and xSuv4-20h2 (B) mRNA expression was detected by RNA *in situ* hybridization at the indicated developmental stages. (C) Total RNA was extracted from animal cap (AC), marginal zone (MZ) and vegetal pole (VP) explants of NF9 embryos; semiquantitative PCR shows relative levels of xSuv4-20h1 and xSuv4-20h2 transcripts in the three explants. ODC was used as loading control, -RT as negative

control. (D) qRT-PCR profiles of xSuv4-20h enzymes. The chart shows the expression of the two enzymes relative to ODC at the indicated developmental stages. (PDF)

Figure S3 Morpholino specificity. 25-mer xSuv4-20h1 morpholino (A) and xSuv4-20h2 morpholino (B) oligonucleotides perfectly target to the start codon of the respective two non-allelic isoforms. Sequence differences between the two morpholinos confer specificity of each oligonucleotide for either xSuv4-20h1 or xSuv4-20h2 mRNA. (C) *In vitro* TNT assay performed as described in Materials and Method section. xSuv4-20h1 and h2 MOs specifically inhibited translation of their cognate templates. (PDF)

Figure S4 Quantification of histone methylation states in xSuv4-20h morphants by MALDI-TOF mass spectrometry. Bulk histones from NF30-33 embryos were isolated and analysed as described in Materials and Methods. (A) H4 peptide 20–23, (B) H3 peptide 9–17 and (C) H3 peptide 27–40 in uninjected, ctrl-MO and double morphant embryos. The values represent mean of three independent experiments; error bars indicate SEM. Star - for technical reasons H4K20me3 mark was quantitated only in some samples (star). (D) Western Blot analysis of uninjected, control morpholino (ctrl-MO) and xSuv4-20h1, h2 morpholino injected embryos using antibodies against H3K27me3 and H3K9me3. PanH3 antibody was used as loading control. (PDF)

Figure S5 Germ layer marker gene expression in xSuv4-20h double morphants. (A) Immuno-histochemistry on ctrl-MO and xSuv4-20h double-morphant tadpoles. Panels show representative cross-sections of neural tubes stained with antibodies against the histone epitopes indicated on top. Inj – injected side. Squares on double-morphant sections represent the cropped pictures shown in Figure 1D. (B-F) RNA *in situ* hybridization analysis of ctrl-MO injected and double morphant embryos at the indicated stages using probes against Chordin, Xnr-3 and Goosoid (B), Krox20 and Otx2 (C), Pax-6 (D), N-CAM (E), FoxD5, Geminin, Zic2, Zic3, Sox3, Sox11 and VegT (F). Chordin, Xnr-3 and Goosoid – dorsal side views; animal pole is on the top. Krox20 – dorsal views, anterior on the left. Otx2 – anterior views, dorsal on the top. Pax-6 – head region; rescued embryos included. N-CAM – dorsal views of stained embryos with the anterior on the left; rescued embryos included. FoxD5, Geminin, Zic2, Zic3 – dorsal views; injected halves are lineage-traced by coinjection of LacZ mRNA and subsequent β -Gal staining (light blue). Sox2, Sox3 – dorsal views, anterior to the left. VegT – internal stain in bisected embryos. (PDF)

Figure S6 Functional SET domains are required for proper Suv4-20h activity. (A) Schematic of *Mus musculus* Suv4-20h1 and h2 SET domain mutations. (B) Immunofluorescence of wild-type and Suv4-20h DKO MEFs transfected with the indicated constructs. (C) Western Blot analysis of NF11.5 uninjected embryos (lane 1), xSuv4-20h1, h2 double morphants (lane 2), active and inactive mouse Suv4-20h1, h2 mRNAs injected embryos (lane 3, 5 respectively) and double morphant embryos coinjected with active or inactive mSuv4-20h1, h2 mRNAs (lane 2, 4 respectively), using antibodies against H4K20 mono-, di- and trimethylation. PanH3 antibody was used as loading control. (D) Anti-myc western blot with the same samples used in B. Asterisks indicate unspecific bands. (PDF)

Figure S7 A functional SET domain is required for morphological and molecular rescue of double-morphants phenotypes. (A–C) Morphological phenotypes of NF30-33 double morphants (A), embryos injected with xSuv4-20h1, h2 morpholinos and active (active rescue, B), or inactive (inactive rescue, C) mouse Suv4-20h1, h2 mRNAs. Embryos were coinjected in one half at two cell stage with Alexa Fluor 488 Dextran as lineage tracer (green channel) to identify the injected side and sort embryos. (D) Penetrance of the eye phenotype in the indicated samples. Displayed are the results from two independent experiments. (E) RNA *In situ* hybridization analysis of NF15 uninjected, double morphants, active and inactive rescue embryos using probes against Oct-25, N-tubulin and MyoD. The pictures show dorsal view of stained embryos, anterior is on the left. (PDF)

Figure S8 *Xenopus laevis* Suv4-20h1 or h2 mRNA overexpression. (A) Overexpression of frog Suv4-20h1 and h2 enzymes causes an upregulation of H4K20me2 and H4K20me3 marks. Bulk histones from uninjected embryos or embryos bilaterally injected with increasing amounts of Suv4-20h1 or h2 mRNAs were isolated at NF11.5 and analysed by Western blot. Pan H3 antibody was used as loading control. (B, C) Morphological phenotypes of NF30-33 embryos injected with xSuv4-20h1 (B) or h2 (C) mRNA. (D) RNA *In situ* hybridization of NF30-33 uninjected embryos (top row) and embryos injected with Suv4-20h1 (middle row) or h2 (bottom row) mRNA using probes against Rx-1 and Pax-6. Pictures show the head of stained embryos. (E) RNA *In situ* hybridization analysis of uninjected embryos (top row) and embryos injected with xSuv4-20h1 (middle row) or h2 (bottom row) mRNA using probes against Ngnr-1a, Delta-like 1, N-tubulin, Xbra, MyoD, Sox17 α and Endodermin. Pictures show dorsal views of stained embryos, anterior is on the left; Xbra pictures show vegetal views of NF11 embryos; MyoD pictures show dorsal views of NF15 embryos, with the head on the left. For Sox-17 α and Endodermin sagittal sections of NF15 embryos were created; pictures show internal view of the injected halves, with anterior on the left. (PDF)

Figure S9 *Mus musculus* Suv4-20h1 or h2 mRNA overexpression. (A) Western Blot analysis of uninjected embryos or embryos injected with *Mus musculus* Suv4-20h1 or h2 mRNAs at different concentrations. Bulk histones from NF11.5 embryos were isolated and analyzed as described in Materials and Methods section. Pan H3 antibody was used as loading control. (B, C) Morphological phenotypes of NF30-33 embryos injected with mouse Suv4-20h1 (B) or h2 (C) mRNA. (D) RNA *In situ* hybridization analysis of uninjected embryos (top row) and embryos injected with mSuv4-20h1 (middle row) or h2 (bottom row) mRNA using probes against Ngnr 1a, Delta-like 1 and Rx-1. Rx-1 pictures show the head of NF30-33 stained embryos. Ngnr 1a (NF12.5) and Delta-like 1 (NF13) pictures show dorsal view of stained embryos, anterior is on the left. (PDF)

Figure S10 Cell proliferation and apoptosis in xSuv4-20h double morphants. (A) Double morphants show increased number of apoptotic cells during neurulation. Top row – immunocytochemistry for active Caspase3 in unilaterally injected embryos (NF15). Middle and bottom rows – RNA *in situ* hybridisation for Delta-like 1 (NF13) and N-tubulin mRNAs (NF15). Pictures show dorsal views, with anterior to the left. (B) Proliferation assay – immunocytochemistry for the mitotic histone modification H3S10P in Ctrl-MO *versus* double morphant embryos. The chart shows a two-fold difference in the number of H3S10P positive cells

on the injected side of double morphants. Data represent mean values of four embryos per condition from two independent experiments; error bars indicate SEM. (PDF)

Figure S11 qRT-PCR analysis. (A) Schematic representation of mRNA purification from NF14-15 embryos for qRT-PCR experiments. (B) qRT-PCR profiles for the indicated genes in Ctrl-MO injected and xSuv4-20h double morphant embryos. Data represent normalized ratios of mRNA levels as means of four independent experiments, error bars indicate SEM. (PDF)

Figure S12 Microarray analysis. (A) Schematic representing mRNA purification from NF 14–15 embryos for microarray experiments. (B) Pie-chart showing number of up- (green) and down- (red) regulated genes. (C) Histogram summarizing the fold expression change of the analysed 9752 active genes. Indicated in red are responder genes (153up, 169 down). (D) Table presenting the 10 most upregulated genes. For each gene, the gene name, symbol, the log fold change (logFC) and the fold change are indicated. (PDF)

Figure S13 Oct-25 and N-tubulin gene expression in *X. tropicalis* Suv4-20h double morphant embryos. RNA *in situ* hybridization analysis of ctrl-MO injected and double morphant embryos at neurula stage (NF14-15) using probes against Oct-25 and N-tubulin. The pictures show dorsal views of the open neural plate with anterior to the left. Dashed line: embryonic midline. (PDF)

Figure S14 Genes analysed by ChIP. (A) Schematic representation of the genes and the amplicons analysed in ChIP experiments. Black boxes: exons; white boxes: untranslated regions; line connecting boxes: introns. For each gene the position of the amplicon(s) used in the experiments is indicated below the gene structure. (B) qRT-PCR profiles for GAPDH and thra genes in Ctrl-MO injected and xSuv4-20h double-morphant embryos. Data represent normalized ratios of mRNA levels as means of five independent experiments, error bars indicate SEM. (PDF)

Figure S15 Regulation of early neural marker genes by Oct-25-VP16 and Oct-25-EnR fusion proteins. (A) qRT-PCR on animal cap (AC) explants cut from uninjected embryos and embryos overexpressing Oct-25-VP16 or Oct-25-EnR mRNAs. The chart shows the relative expression of the indicated genes compared to H4 gene levels. Data represent normalized ratios of mRNA levels as means of three or four independent experiments, error bars indicate SEM. (B) *In situ* hybridization on uninjected AC or explants overexpressing Oct-25-VP16/EnR for Zic1 (upper row, 20 \times magnification) and Zic3 genes (lower row, 50 \times magnification). (PDF)

Figure S16 Oct-25 binding sites on Zic1, Zic3 and Sox2 genes. Oct-25 hypothetical binding sites on Zic1, Zic3 and Sox2 have been identified as described in the Material and Methods section. The schematic representation of the genes shows: black boxes: exons; white boxes: untranslated regions; line connecting boxes: introns. For each gene the position of the binding sites, the identified sequence and the similarity to the published weight matrix (weight) are indicated in the underneath table. For Zic3 the six highest identified sequences are shown. (PDF)

Figure S17 ES cell analysis. (A) Morphological appearance of differentiated day 2 embryoid bodies: B4-2 and B7-1 cell lines

formed smaller bodies than those of the wild-type line wt26. (B) GSES-1 morphology and *in vitro* differentiation at the indicated days. Scale bar: 100 μ m. (C) FACS profiles of GSES-1 cell line before (day 0) and during (day 6) differentiation. Red graphs: fluorescence of non-specific isotype control; black graphs: Oct4 protein levels in GSES-1 cells; green dashed graphs: Oct4 protein levels in mutant B7-1 cell line. (D) qRT-PCR profiles for the indicated genes in wild-type (wt) and Suv4-20h DKO cell lines at differentiation day 6. FoxA2 and Gata4 expression levels are misregulated in Suv4-20h DKO cells upon differentiation. (PDF)

Table S1 List of oligonucleotide sequences used in this study. (PDF)

Text S1 Supporting information on experimental procedures. This file contains additional information on statistical analysis, extraction of Myc-tagged fusion protein from embryos, qRT-PCR samples preparation, Vibratome sections of Oct-25 stained embryos, Immunostaining, Immunofluorescence microscopy of MEF cells and ChIP (chromatin immunoprecipitation) analysis. (DOC)

References

- Niwa H (2007) How is pluripotency determined and maintained? *Development* 134: 635–646.
- Meshorer E, Yellajoshula D, George E, Scambler PJ, Brown DT, et al. (2006) Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev Cell* 10: 105–116.
- Hemberger M, Dean W, Reik W (2009) Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington's canal. *Nat Rev Mol Cell Bio* 10: 526–537.
- Saladi SV, de la Serna IL (2010) ATP dependent chromatin remodeling enzymes in embryonic stem cells. *Stem Cell Rev* 6: 62–73.
- Singhal N, Graumann J, Wu G, Arauzo-Bravo MJ, Han DW, et al. (2010) Chromatin-Remodeling Components of the BAF Complex Facilitate Reprogramming. *Cell* 141: 943–955.
- Banaszynski LA, Allis CD, Lewis PW (2010) Histone variants in metazoan development. *Dev Cell* 19: 662–674.
- Lee BM, Mahadevan LC (2009) Stability of histone modifications across mammalian genomes: implications for 'epigenetic' marking. *J Cell Biochem* 108: 22–34.
- Meshorer E, Misteli T (2006) Chromatin in pluripotent embryonic stem cells and differentiation. *Nat Rev Mol Cell Bio* 7: 540–546.
- Biron VL, McManus KJ, Hu N, Hendzel MJ, Underhill DA (2004) Distinct dynamics and distribution of histone methyl-lysine derivatives in mouse development. *Dev Biol* 276: 337–351.
- Hashimshony T, Zhang J, Keshet I, Bustin M, Cedar H (2003) The role of DNA methylation in setting up chromatin structure during development. *Nat Genet* 34: 187–192.
- Lee JH, Hart SR, Skalik DG (2004) Histone deacetylase activity is required for embryonic stem cell differentiation. *Genesis* 38: 32–38.
- Akkers RC, van Heeringen SJ, Jacobi UG, Janssen-Megens EM, Francoijs KJ, et al. (2009) A hierarchy of H3K4me3 and H3K27me3 acquisition in spatial gene regulation in *Xenopus* embryos. *Dev Cell* 17: 425–434.
- Schneider TD, Arteaga-Salas JM, Mentale E, David R, Nicetto D, et al. (2011) Stage-specific histone modification profiles reveal global transitions in the *Xenopus* embryonic epigenome. *PLoS ONE* 6: e22548. doi:10.1371/journal.pone.0022548
- Vastenhouw NL, Zhang Y, Woods IG, Imam F, Regev A, et al. (2010) Chromatin signature of embryonic pluripotency is established during genome activation. *Nature* 464: 922–926.
- Bhaumik SR, Smith E, Shilatifard A (2007) Covalent modifications of histones during development and disease pathogenesis. *Nat Struct Mol Biol* 14: 1008–1016.
- Dambacher S, Hahn M, Schotta G (2010) Epigenetic regulation of development by histone lysine methylation. *Hereditas* 105: 24–37.
- Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, et al. (2006) Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* 441: 349–353.
- Cao R, Zhang Y (2004) The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. *Curr Opin Genet Dev* 14: 155–164.
- Christophersen NS, Helin K (2010) Epigenetic control of embryonic stem cell fate. *J Exp Med* 207: 2287–2295.
- Rea S, Eisenhaber F, O'Carroll D, Strahl BD, Sun ZW, et al. (2000) Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* 406: 593–599.
- Schotta G, Lachner M, Sarma K, Ebert A, Sengupta R, et al. (2004) A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Gene Dev* 18: 1251–1262.
- Tachibana M, Sugimoto K, Nozaki M, Ueda J, Ohta T, et al. (2002) G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Gene Dev* 16: 1779–1791.
- Yang L, Xia L, Wu DY, Wang H, Chansky HA, et al. (2002) Molecular cloning of ESET, a novel histone H3-specific methyltransferase that interacts with ERG transcription factor. *Oncogene* 21: 148–152.
- Schotta G, Sengupta R, Kubicek S, Malin S, Kauer M, et al. (2008) A chromatin-wide transition to H4K20 monomethylation impairs genome integrity and programmed DNA rearrangements in the mouse. *Gene Dev* 22: 2048–2061.
- Dillon SC, Zhang X, Trievel RC, Cheng X (2005) The SET-domain protein superfamily: protein lysine methyltransferases. *Genome Biol* 6: 227.
- Kwon T, Chang JH, Kwak E, Lee CW, Joachimiak A, et al. (2003) Mechanism of histone lysine methyl transfer revealed by the structure of SET7/9-AdoMet. *EMBO J* 22: 292–303.
- Rogers CD, Moody SA, Casey ES (2009) Neural induction and factors that stabilize a neural fate. *Birth defects research Part C, Embryo today : reviews* 87: 249–262.
- Harris WA, Hartenstein V (1991) Neuronal determination without cell division in *Xenopus* embryos. *Neuron* 6: 499–515.
- Hinkley CS, Martin JF, Leibham D, Perry M (1992) Sequential expression of multiple POU proteins during amphibian early development. *Mol Cell Biol* 12: 638–649.
- Morrison GM, Brickman JM (2006) Conserved roles for Oct4 homologues in maintaining multipotency during early vertebrate development. *Development* 133: 2011–2022.
- Cao Y, Knochel S, Donow C, Miethe J, Kaufmann E, et al. (2004) The POU factor Oct-25 regulates the *Xvent-2B* gene and counteracts terminal differentiation in *Xenopus* embryos. *J Biol Chem* 279: 43735–43743.
- Cao Y, Siegel D, Oswald F, Knochel W (2008) Oct25 represses transcription of nodal/activin target genes by interaction with signal transducers during *Xenopus* gastrulation. *J Biol Chem* 283: 34168–34177.
- Snir M, Ofir R, Elias S, Frank D (2006) *Xenopus laevis* POU91 protein, an Oct3/4 homologue, regulates competence transitions from mesoderm to neural cell fates. *EMBO J* 25: 3664–3674.
- Takebayashi-Suzuki K, Arita N, Murasaki E, Suzuki A (2007) The *Xenopus* POU class V transcription factor XOct-25 inhibits ectodermal competence to respond to bone morphogenetic protein-mediated embryonic induction. *Mech Develop* 124: 840–855.
- Hellsten U, Harland RM, Gilchrist MJ, Hendrix D, Jurka J, et al. (2010) The genome of the Western clawed frog *Xenopus tropicalis*. *Science* 328: 633–636.
- Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, et al. (2007) High-resolution profiling of histone methylations in the human genome. *Cell* 129: 823–837.
- Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, et al. (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448: 553–560.

Video S1 Videorecording of wild-type ES cells at day 14 of differentiation. The video shows several areals of autonomously beating cardiomyocytes. (M4V)

Video S2 Videorecording of Suv4-20h DN ES cells at day 14 of differentiation. In contrast to differentiated wild-type ES cells, spontaneous contractions of cardiomyocytes are not observed. (M4V)

Acknowledgments

We thank Tomas Pieler, Edgard Pera, and Herbert Steinbeisser for *Xenopus* probes; Ulrich Welsch and Astrid Sulz for paraffin sections; Thomas Hollemann for technical advice on *xenopus* sections; Alessandra Pasquarella for help with FACS analysis; Barbara Hölscher, Christiane Gross, and Alexander Nuber for expert technical assistance; Herbert Steinbeisser for comments on the manuscript.

Author Contributions

Conceived and designed the experiments: DN RAWR GS RD. Performed the experiments: DN MH TDS JJ. Analyzed the data: DN MH JJ TDS TS RAWR GS RD. Contributed reagents/materials/analysis tools: GS RD JJ TDS TS. Wrote the paper: DN RAWR.

38. Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, et al. (2005) Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 122: 947–956.
39. Munoz-Sanjuan I, Brivanlou AH (2002) Neural induction, the default model and embryonic stem cells. *Nat Rev Neurosci* 3: 271–280.
40. Cao Y, Siegel D, Knochel W (2006) *Xenopus* POU factors of subclass V inhibit activin/nodal signaling during gastrulation. *Mech Develop* 123: 614–625.
41. Teo AK, Arnold SJ, Trotter MW, Brown S, Ang LT, et al. (2011) Pluripotency factors regulate definitive endoderm specification through *omesodermin*. *Gene Dev* 25: 238–250.
42. Thomson M, Liu SJ, Zou LN, Smith Z, Meissner A, et al. (2011) Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. *Cell* 145: 875–889.
43. David R, Brenner C, Stieber J, Schwarz F, Brunner S, et al. (2008) *MesP1* drives vertebrate cardiovascular differentiation through *Dkk-1*-mediated blockade of Wnt-signalling. *Nature Cell Biol* 10: 338–345.
44. Kapoor-Vazirani P, Kagey JD, Vertino PM (2011) SUV420H2-mediated H4K20 trimethylation enforces RNA polymerase II promoter-proximal pausing by blocking hMOF-dependent H4K16 acetylation. *Mol Cell Biol* 31: 1594–1609.
45. Magklara A, Yen A, Colquitt BM, Clowney EJ, Allen W, et al. (2011) An epigenetic signature for monoallelic olfactory receptor expression. *Cell* 145: 555–570.
46. Zhang Y, Cooke M, Panjwani S, Cao K, Krauth B, et al. (2012) Histone h1 depletion impairs embryonic stem cell differentiation. *PLoS Genet* 8: e1002691. doi:10.1371/journal.pgen.1002691
47. Li Z, Nie F, Wang S, Li L (2011) Histone H4 Lys 20 monomethylation by histone methylase SET8 mediates Wnt target gene activation. *P Natl Acad Sci USA* 108: 3116–3123.
48. Hollemann T, Schuh R, Pieler T, Stick R (1996) *Xenopus* Xsal-1, a vertebrate homolog of the region specific homeotic gene *spalt* of *Drosophila*. *Mech of Develop* 55: 19–32.
49. Blythe SA, Reid CD, Kessler DS, Klein PS (2009) Chromatin immunoprecipitation in early *Xenopus laevis* embryos. *Dev Dynam* 238: 1422–1432.

[illegible][illegible]

mouse Suv4-20h1
460 470 480 490 500 510 520 530 540 550 560 570 580 590 600
S K I R L M P C R L D Q S A S K R L E G S L V L N E K P V L Y N L I P K K E R E P G L A V G S G C G L T R H A R A E R L N H N G R G A S G D L P C Y T T R S L R T A T G L E T T O I L E R S P L D O Y A N G L E P C P S G Q Q T P E V L E A P E L I R E E L
S K R H K . . . Q K A S K . E L E L V N E K P V L Y N L I K A N D R E G G C L T R H A R A E . . . N . G A S G . S C Y T T R S L R N F S K S D I S E A N K L S H O V S N O Y O I R E E L E T . R E E S
xampus Suv4-20h1










mouse Suv4-20h1
xenopus Suv4-20h1

Q E C P N D S C L S R K R E P T H D L A R K R E H S F P G L T G T D P L G C P S T R A P Y S I T D S A S P V C G S L T Y T P S F R A S F R T Q S L
K R F E N N F T T S K R K G I C S K E D S K H A L L D G S H C D G E G I Q T S P D K S F E S A D E P I V S E I T A K R I A K K R R I A D A L L E N S S G I P K I L I R R R D S S K I N E . E G S S A I S I R K I
N S K Q . K . K . E D S K D H . L E D L S H D G E . I T A K R I A K K R R I A D A L L E N S S G I P K I L I R R R D S S K I N E . E G S S A I S I R K I

[illegible][illegible]

xenopus Suv4-20i2
 T L A F P D N D S V M Y T A K A S A Q L E L G P A A F I N H D C R N C N E F P T E N T A C K A V L E I K T T E E I F C E Y G A S F E E H N B M C E C T C E R N G D G F A A Q O M T E Q V S - S L E K Y L R E T D G R K A R S S E S A C P S P Q L T T K G S L R L S T A L K A T I
 L L R G . N D P S . M S T R K . A Q L W L G P A A F I N H D C . P C K C F Y P . G N T A C K A V L R . I G . E . T C F Y G . F P G E K N E C E C T C E R K G . G A F . Q E . P L . K V L R E T G R . R L S . L S T R
 310 320 330 340 350 360 370 380 390 400 410 420 430 440 450

xenopus Suv4-20m2
 P S R R R R G A F F R R R V K F A S S S Y F A S H L H I L P K P V E L I L A G T V I L H V K C K C K C Q A S P S S H R Q C C L E P E L S L R R E D I S E R L E R L R S C P G S C P P L I I T A V G C A N K D C Q T E A G K S N E Q I T A E L C H S L S P V P
 A G A F F R R R V K F A S S S Y F A S S M . P L . . . P V R D V R I I L H V K C K C K C Q A S P S S H R Q C C L E P E L S L R R E D I S E R L E R L R S C P . S P P . . . T A V G C A N K D C Q T E A G K S N E Q I T A E L C H S L S P V P

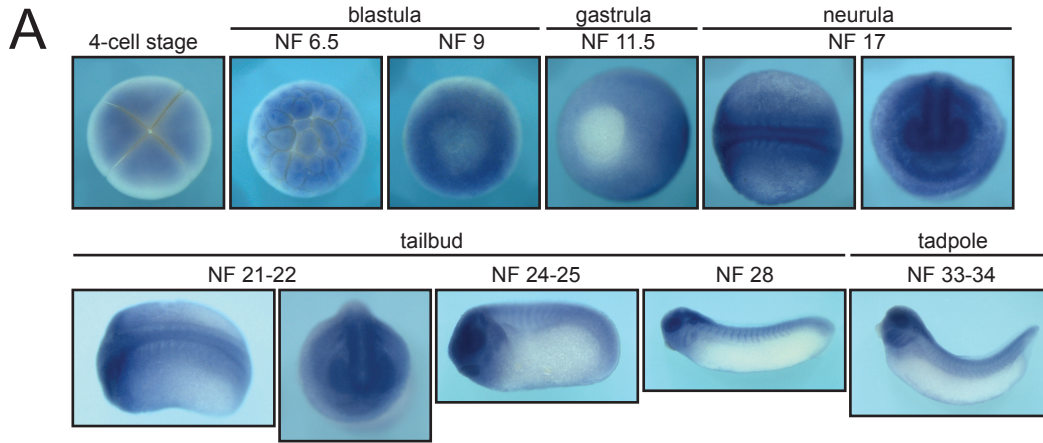
AGIOPLOS SOW4-ZOLV, *H. 2b*         
 610 620 630 640 650 660 670 680 690 700 710 720 730 740 750
 NSFEEHITSL QPESFSESENNLECEASPEPVI NVNVTFFHNGOITSDPHAKLQTCITHTT EA... FQSPQSSSEAEKQIIPNNKSHGEPV SBDHLVANTAEI QSNASFFNTTFCALNOSLEK

mouse Suva-2012 $\frac{P}{D}I \frac{N}{S} \frac{G}{E} \frac{P}{D}P$. . . $\frac{780}{P}$

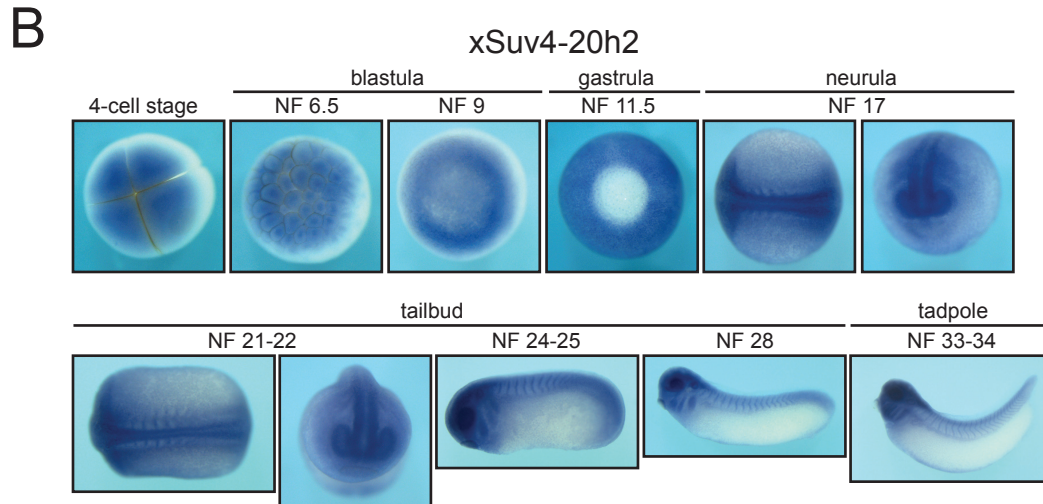
L D S · T S

mSuv4-20h1	98%	83%	-	75%
------------	-----	-----	---	-----

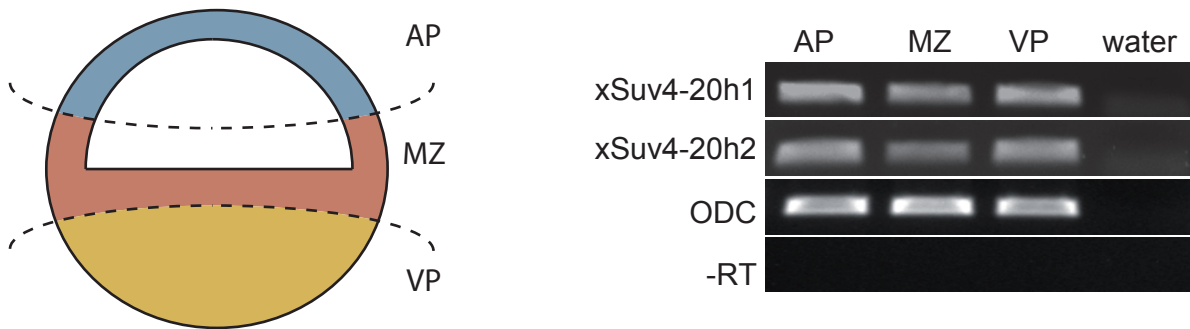
xSuv4-20h1



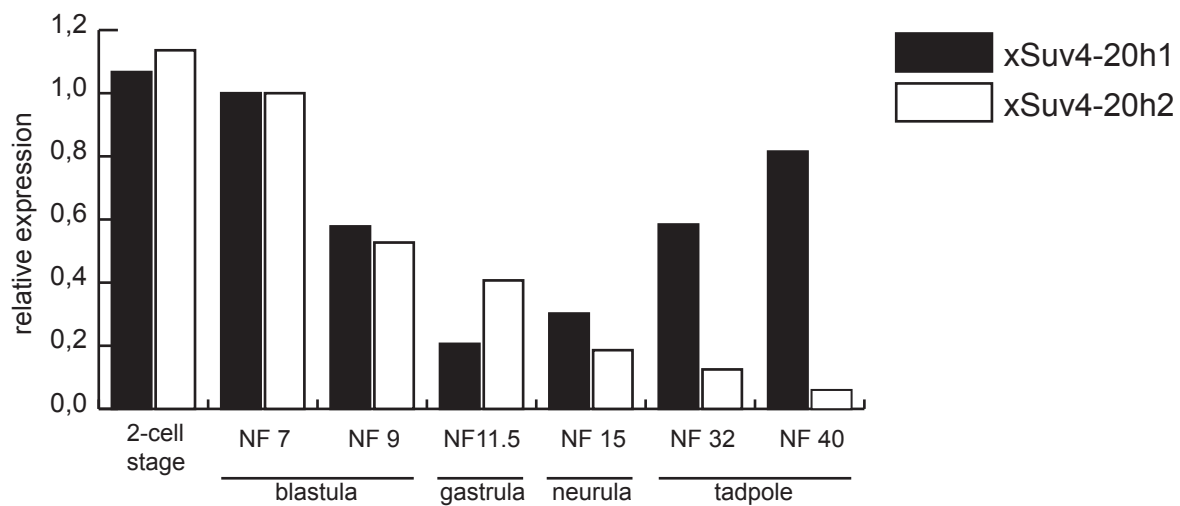
xSuv4-20h2



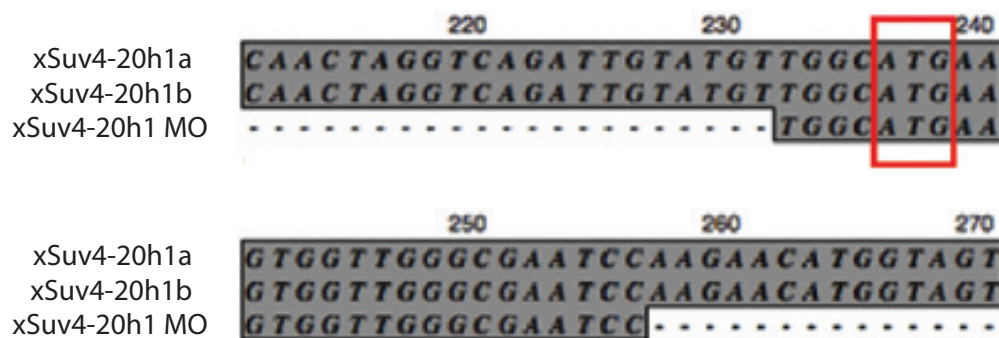
C



D



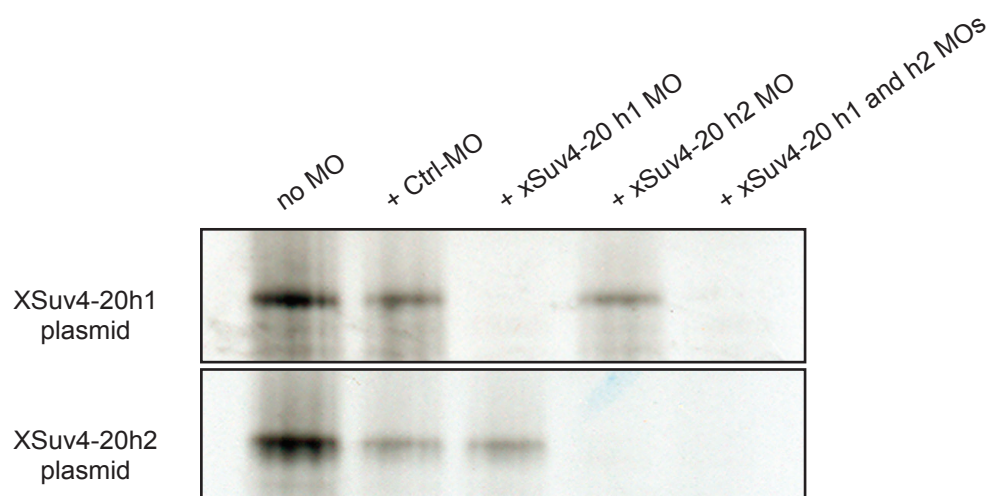
A

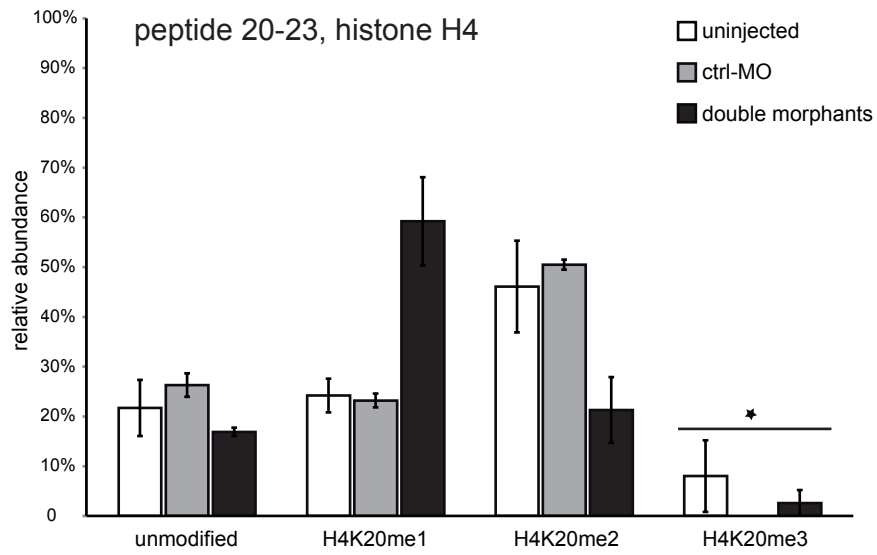
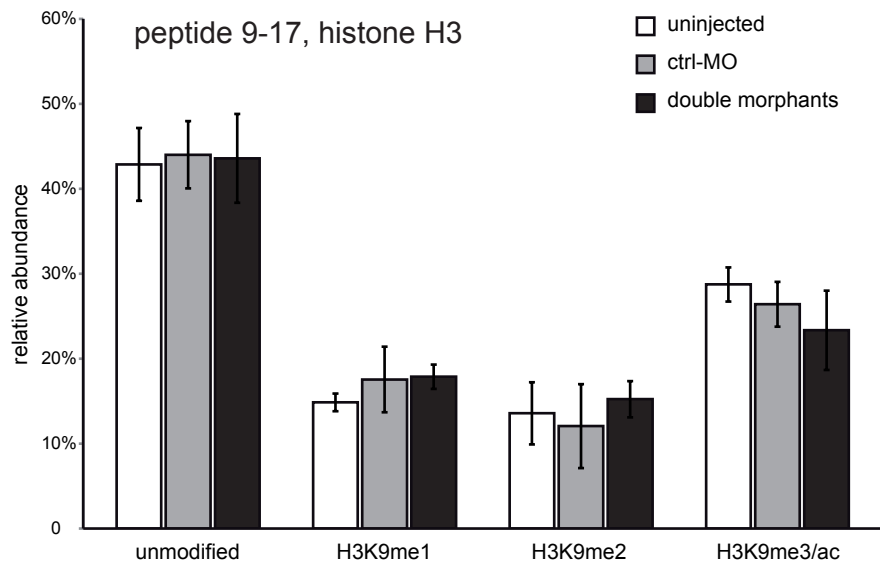
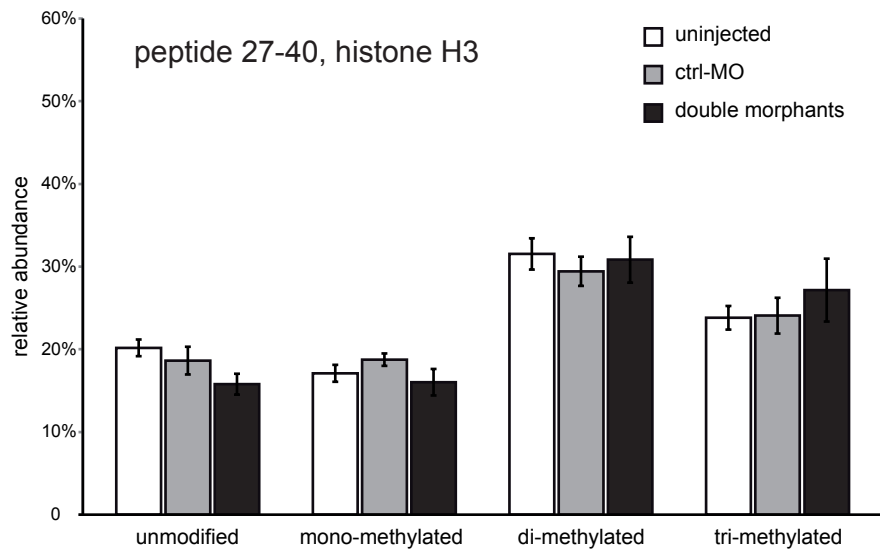
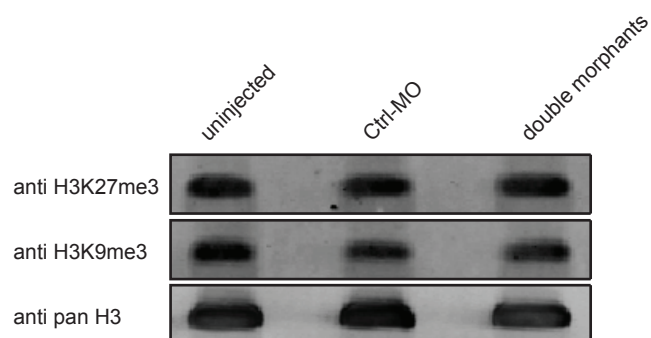


B

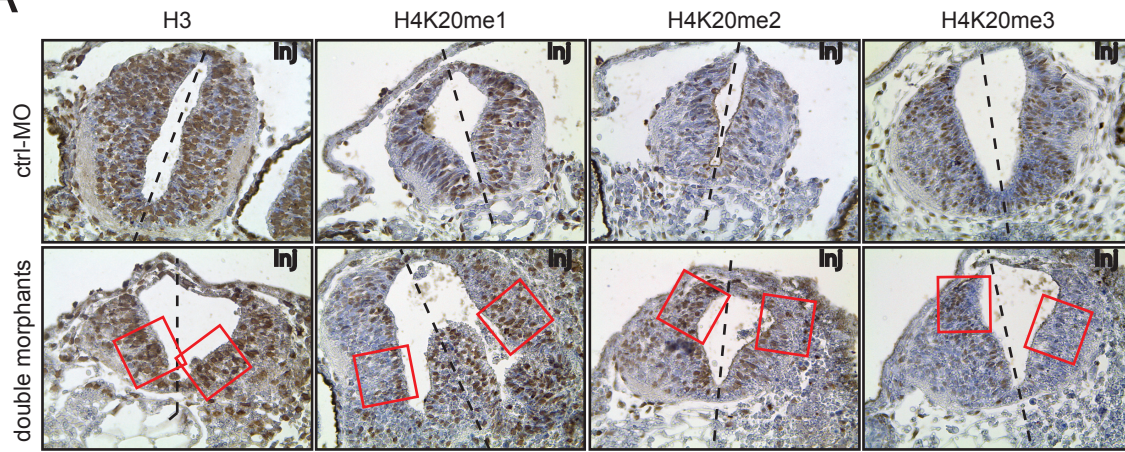


C

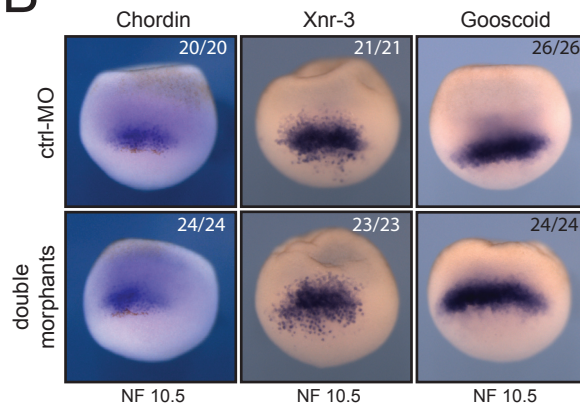


A**B****C****D**

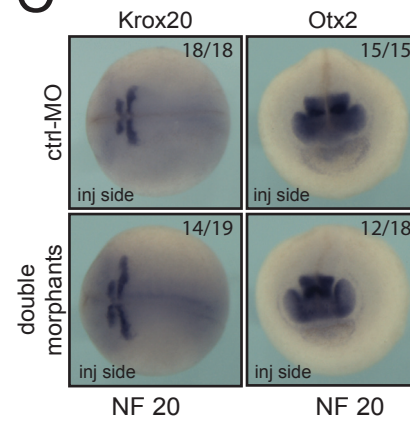
A



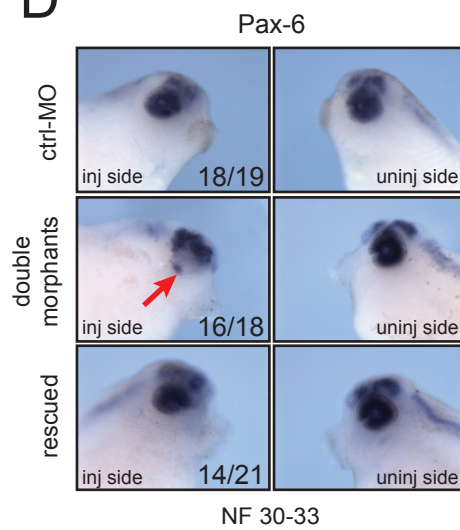
B



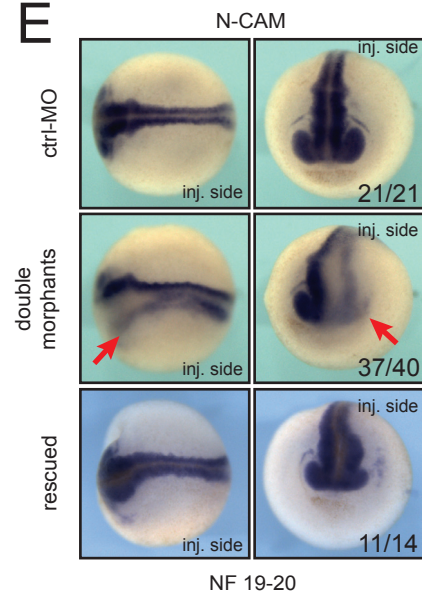
C



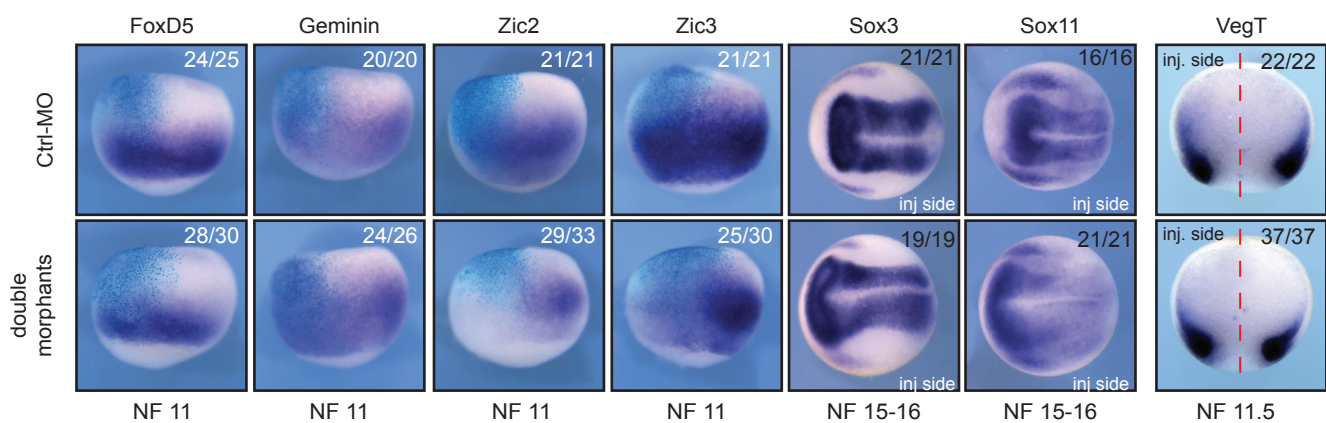
D



E

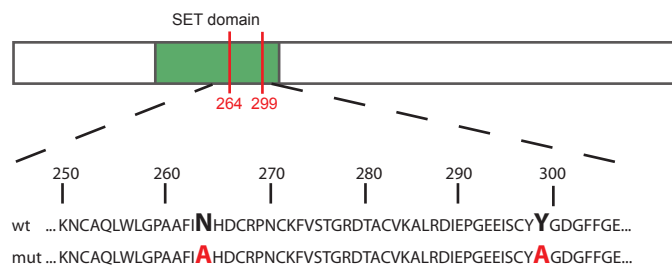


F

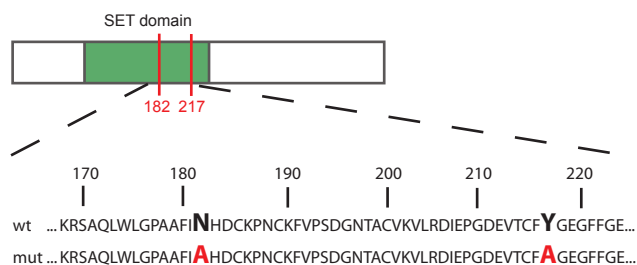


A

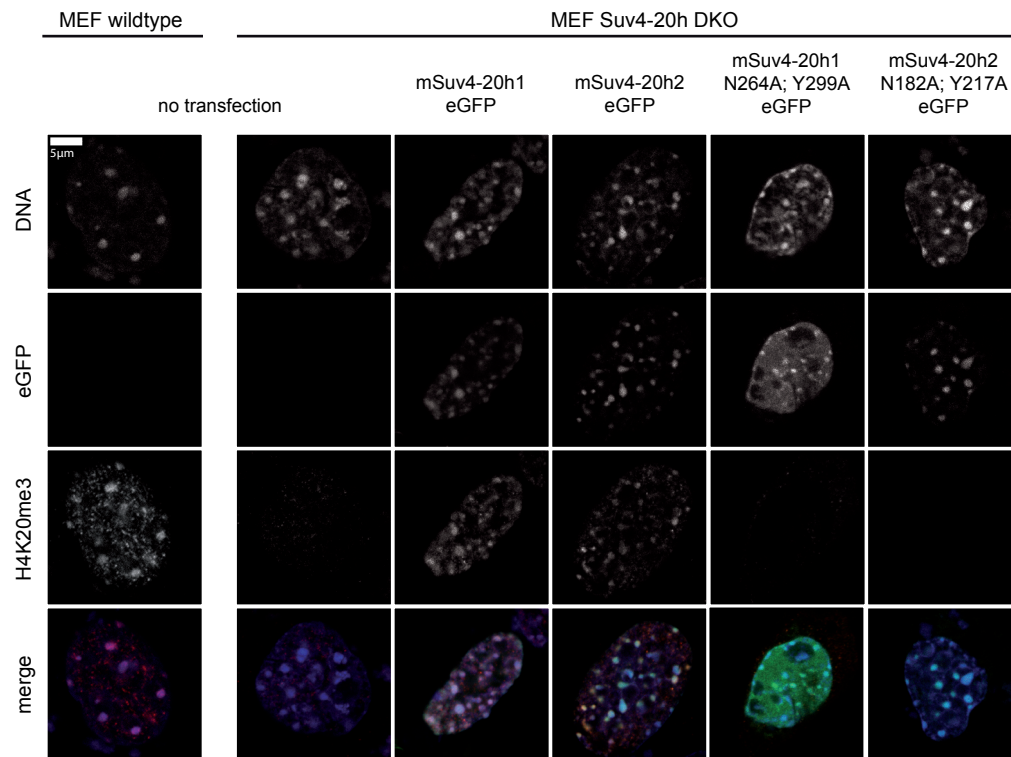
mouse Suv4-20 h1



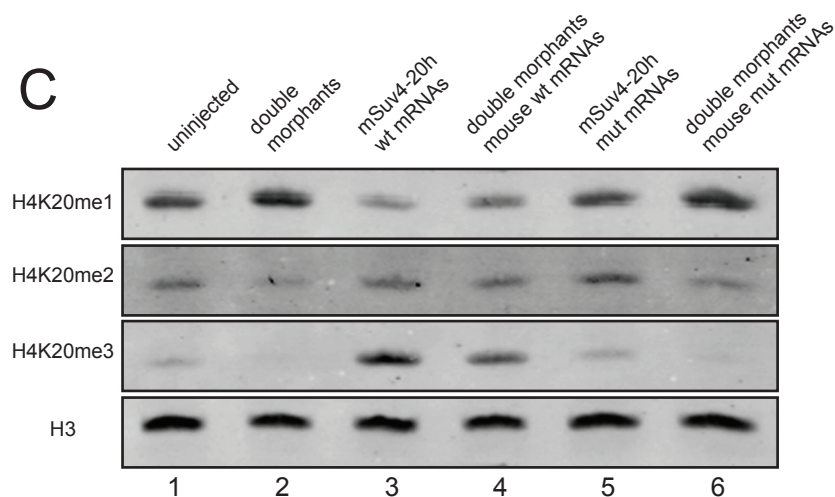
mouse Suv4-20 h2



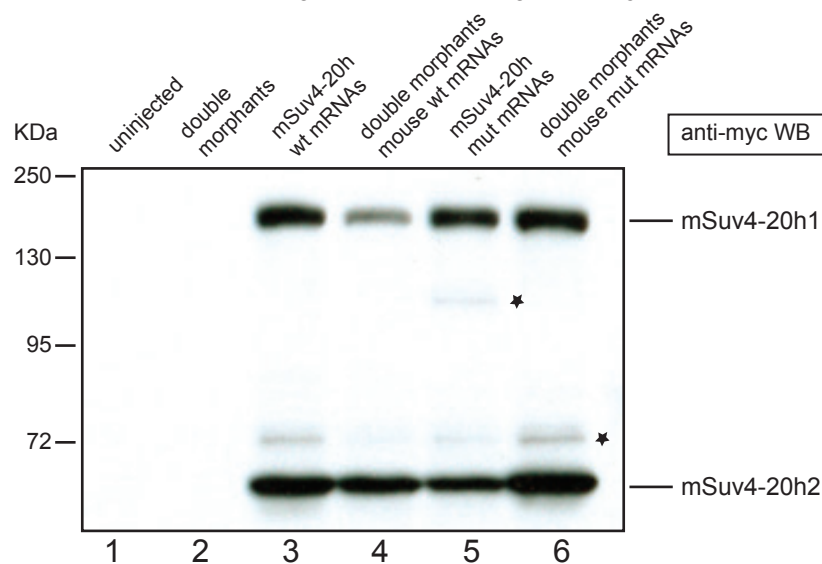
B

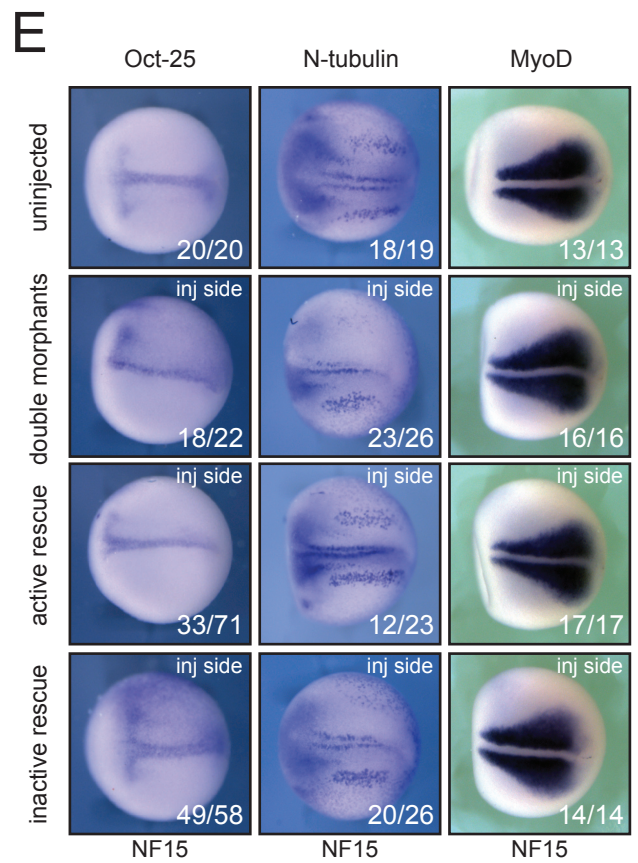
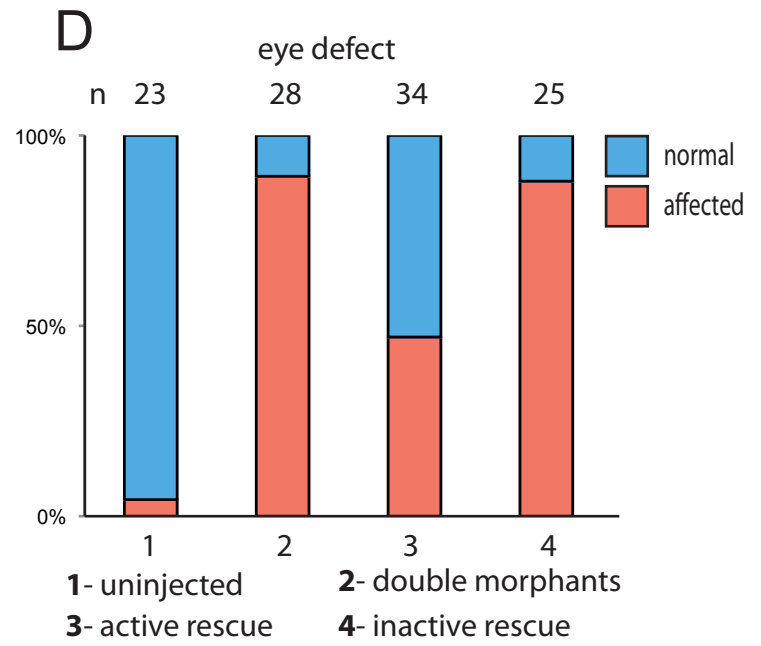
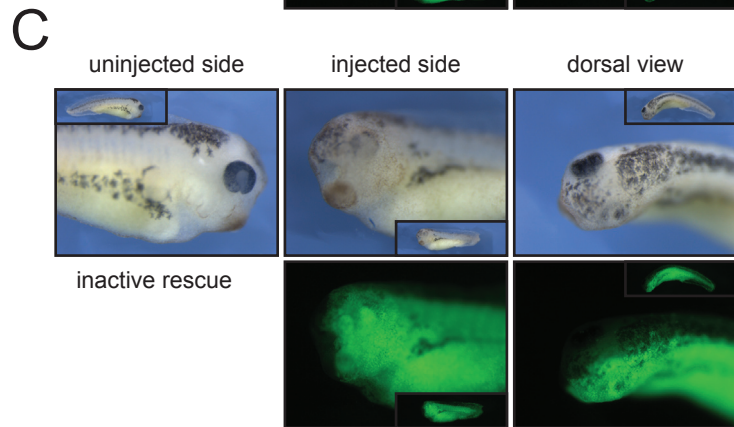
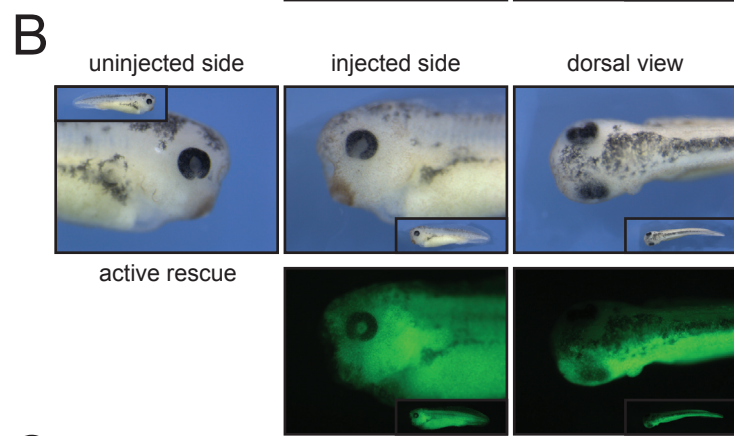
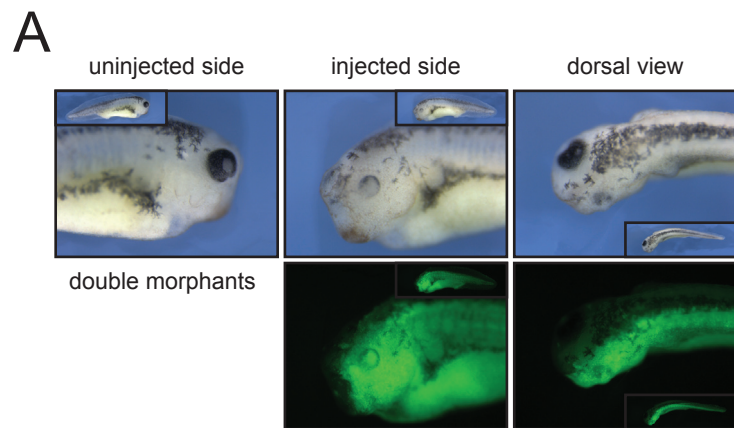


C

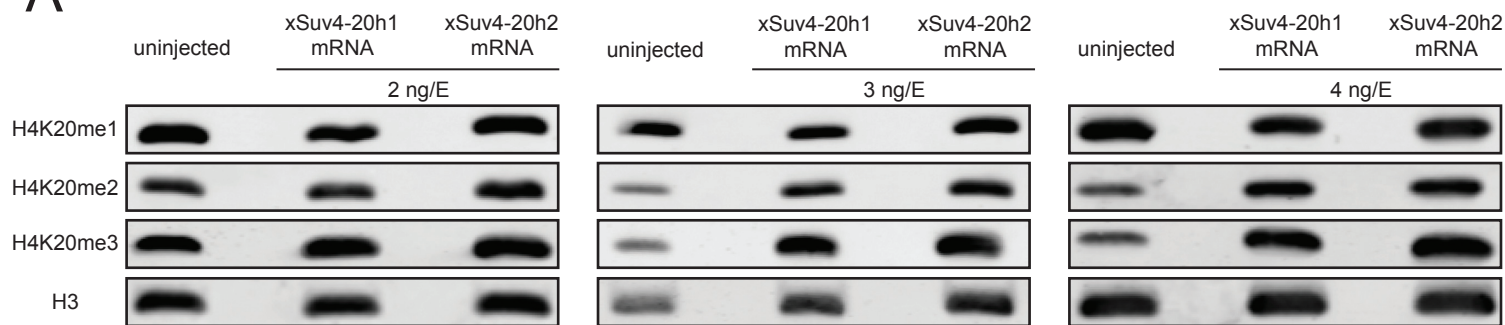


D

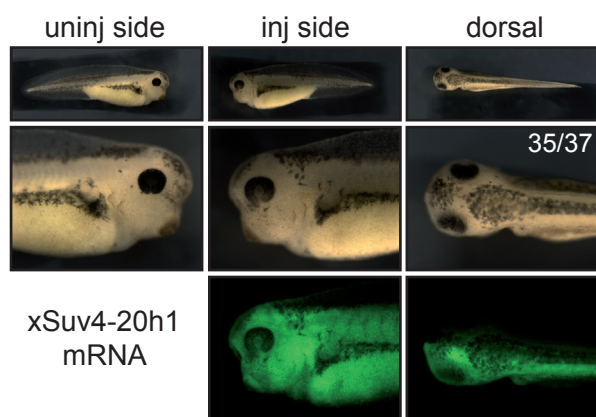




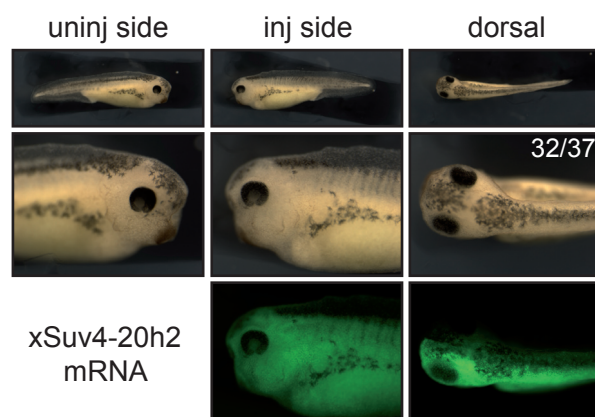
A



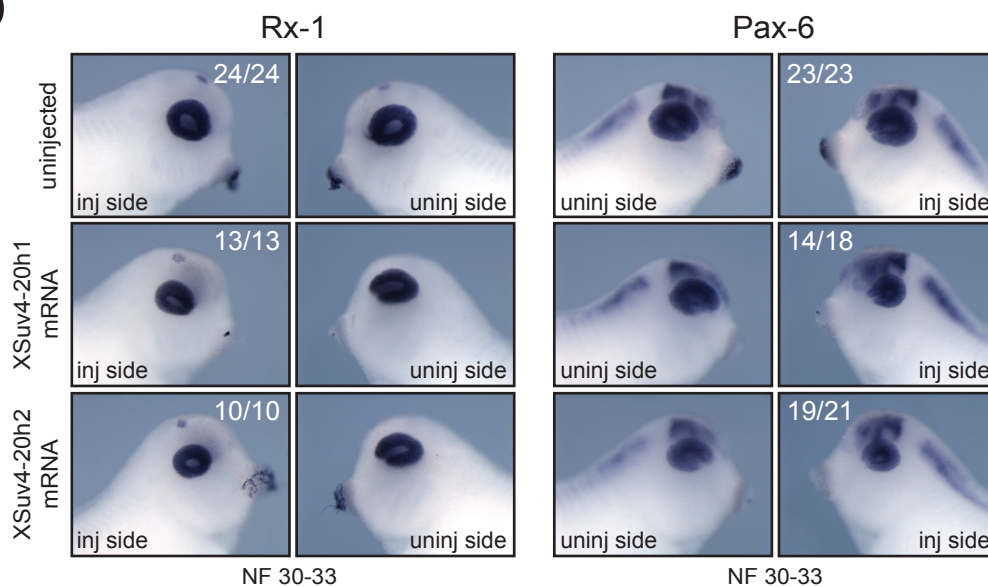
B



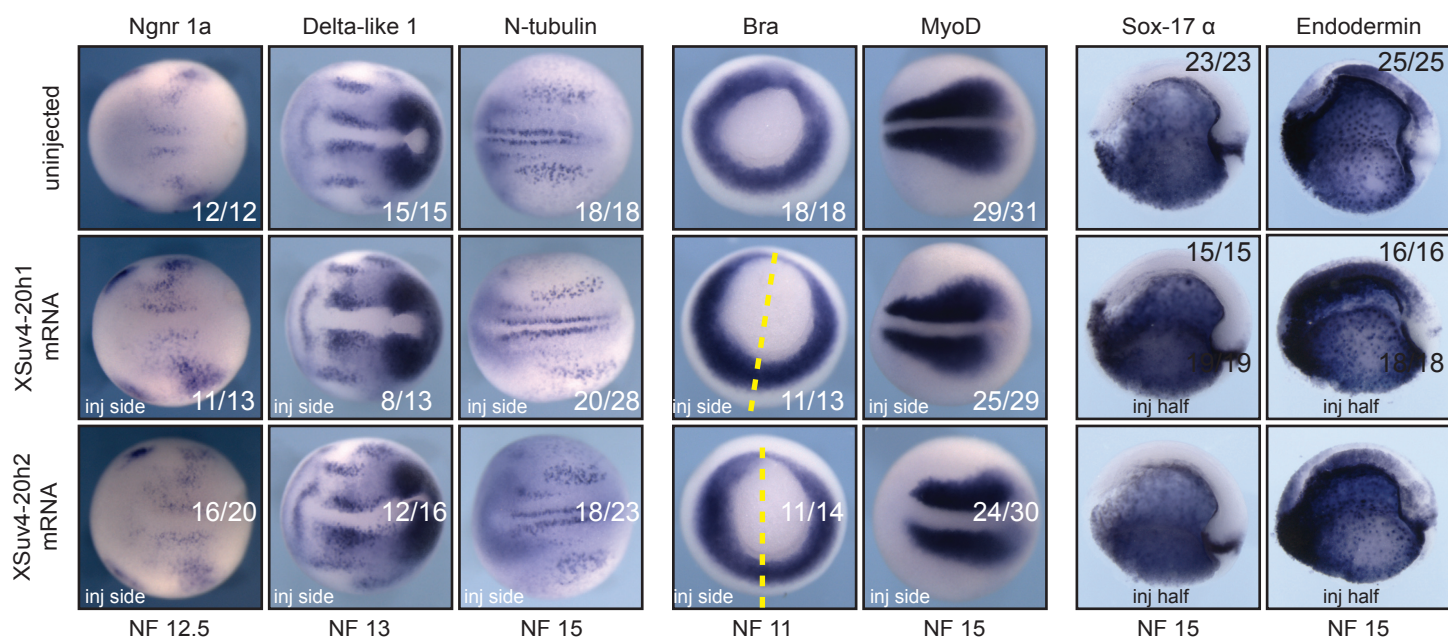
C

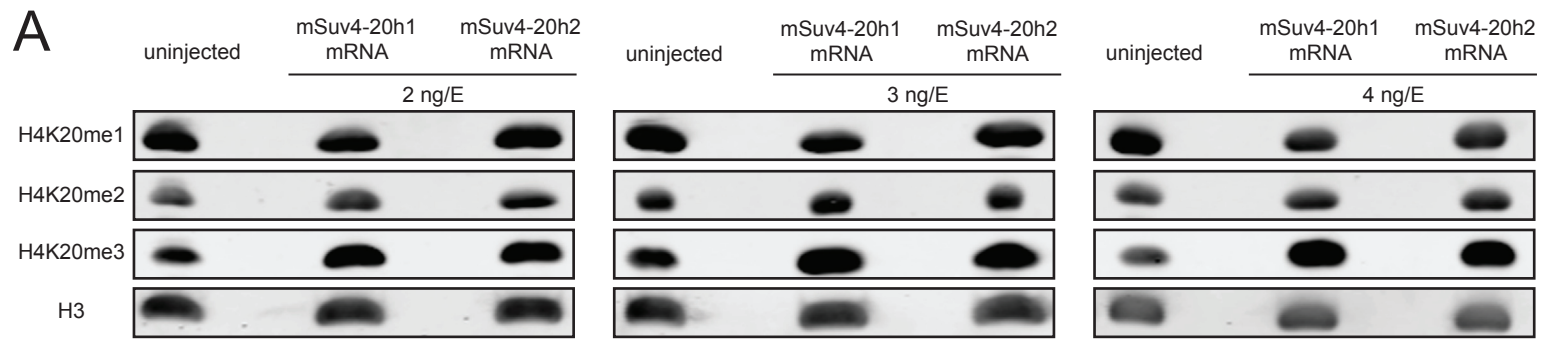


D

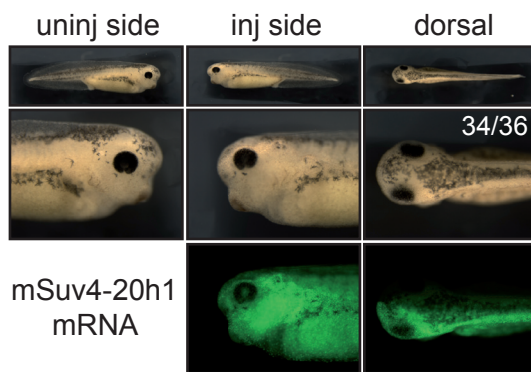


E

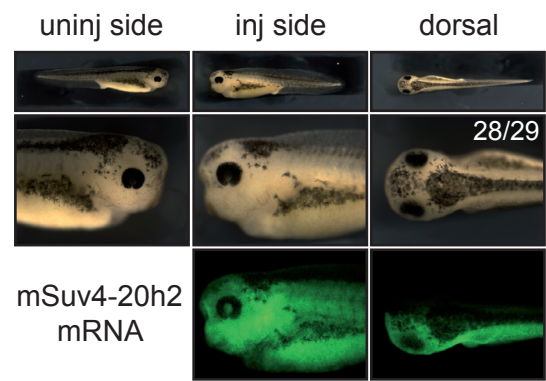




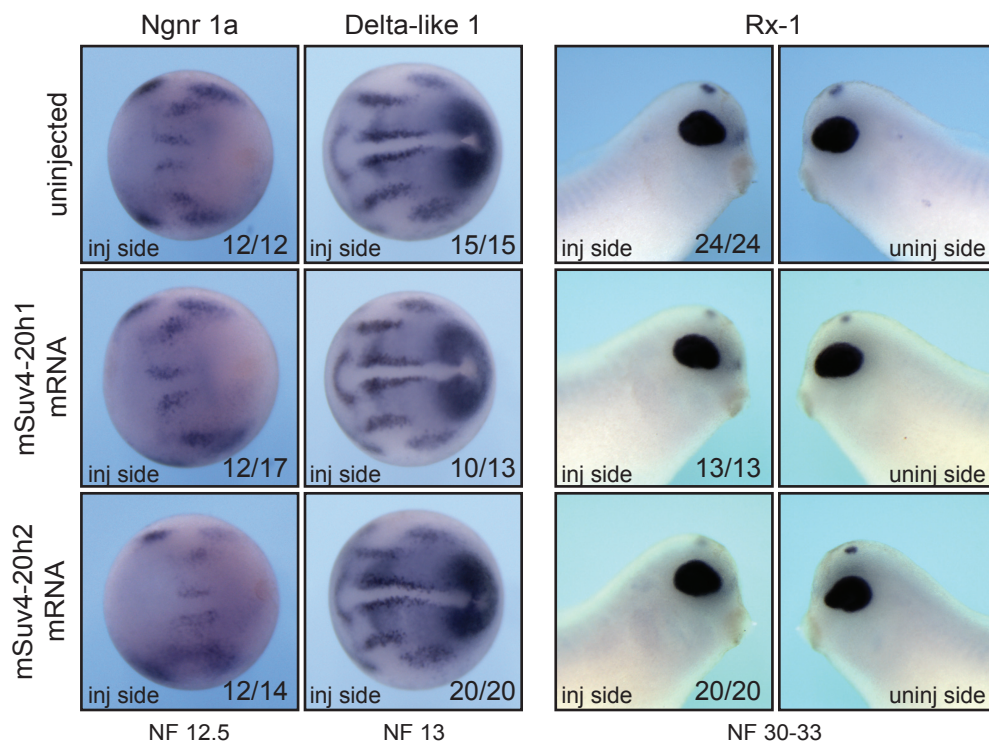
B

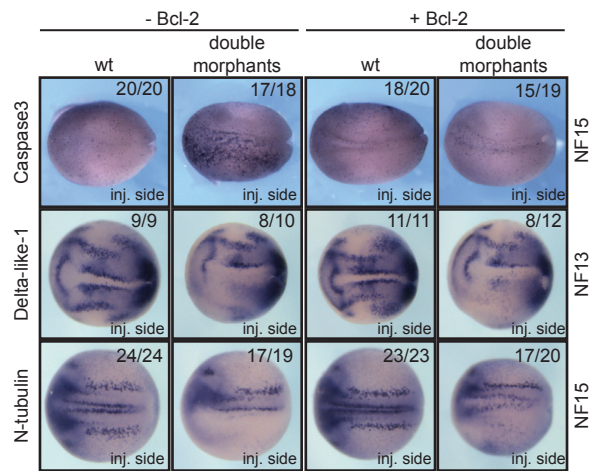
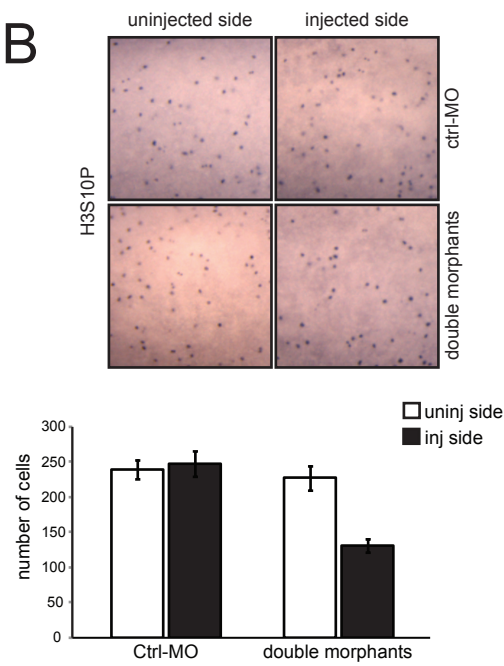


C

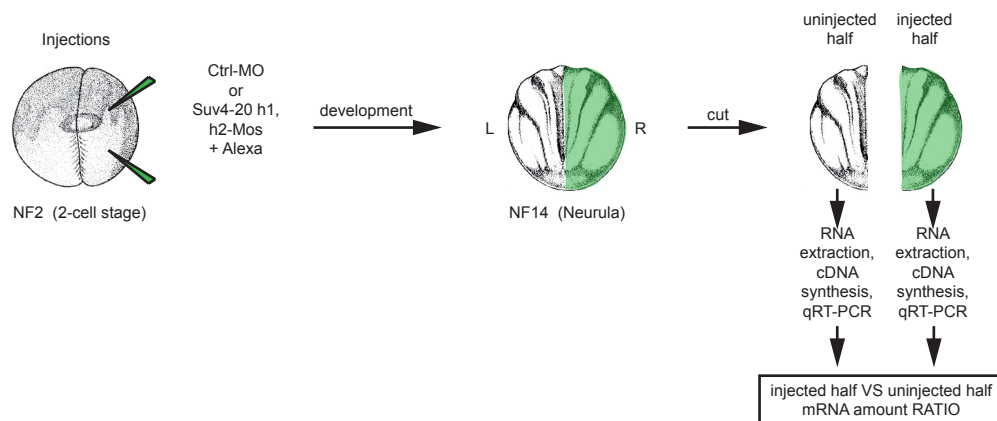


D

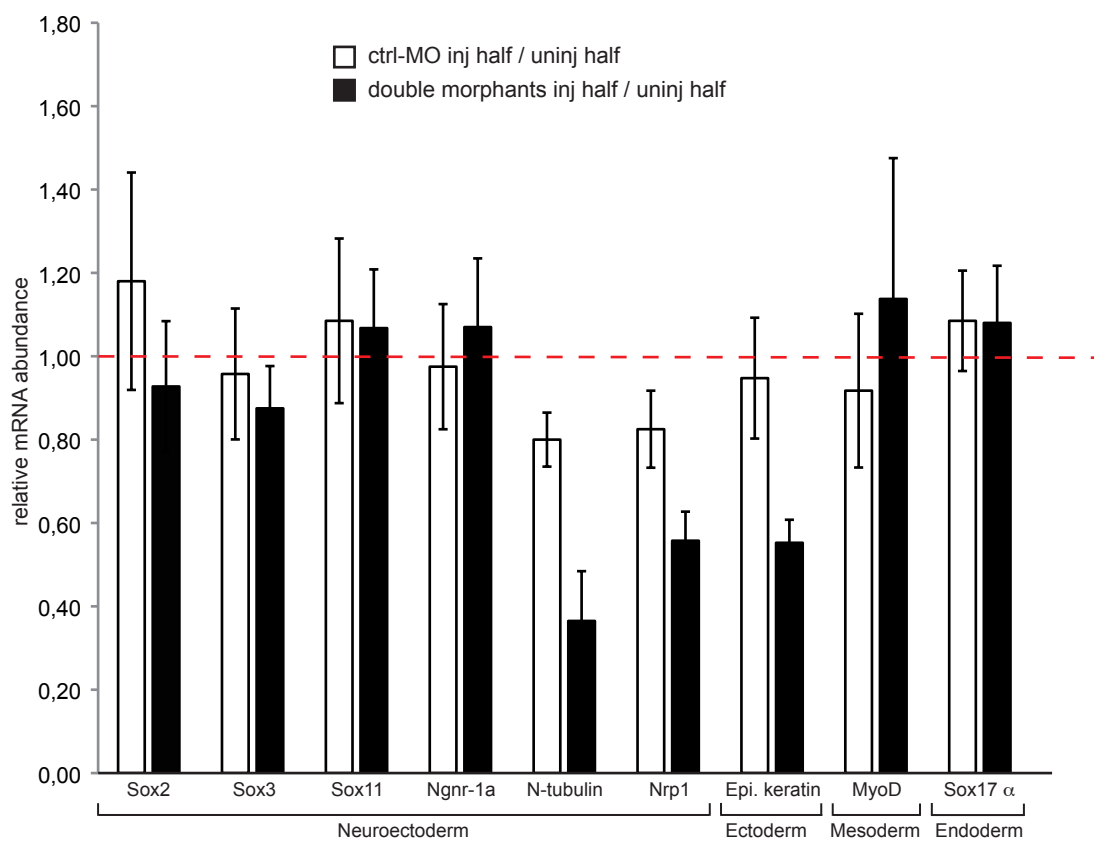


A**B**

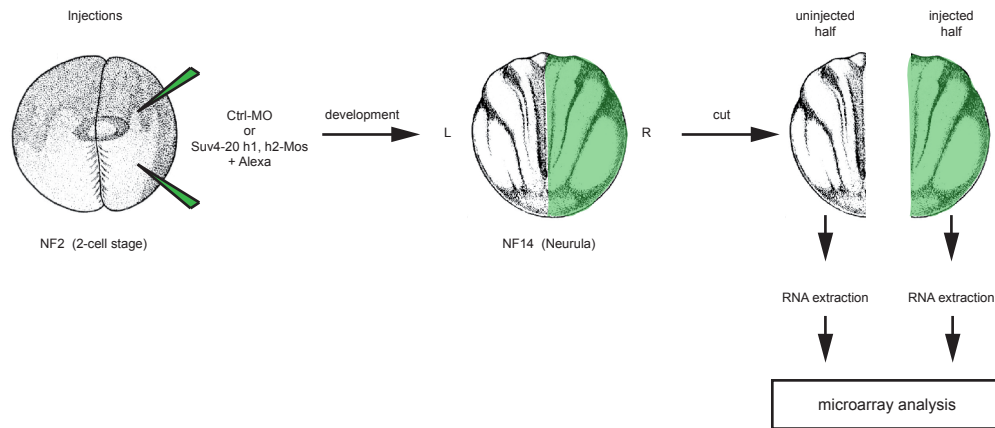
A



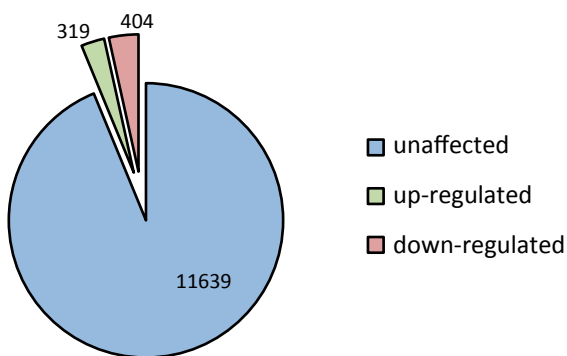
B



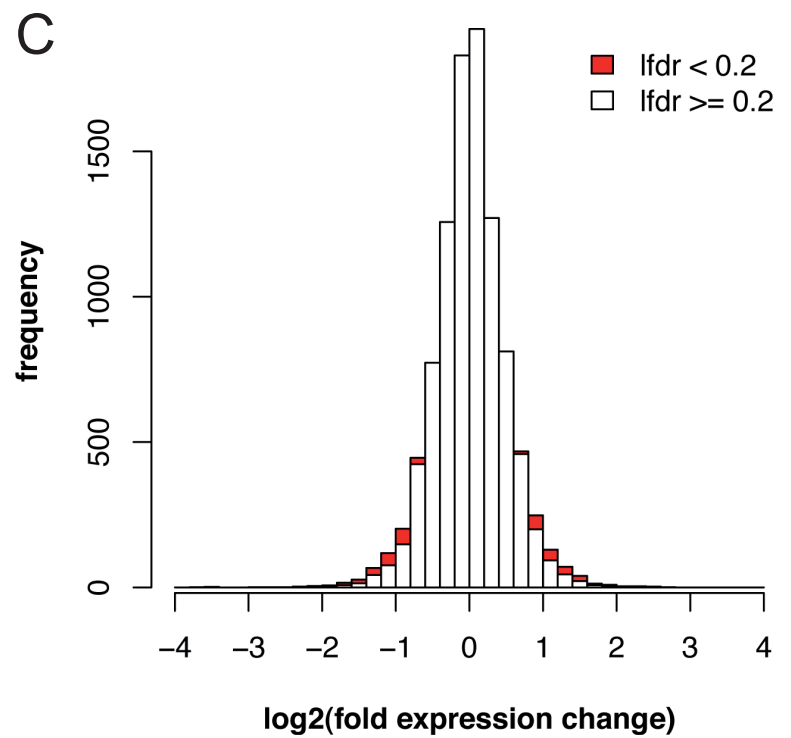
A



B

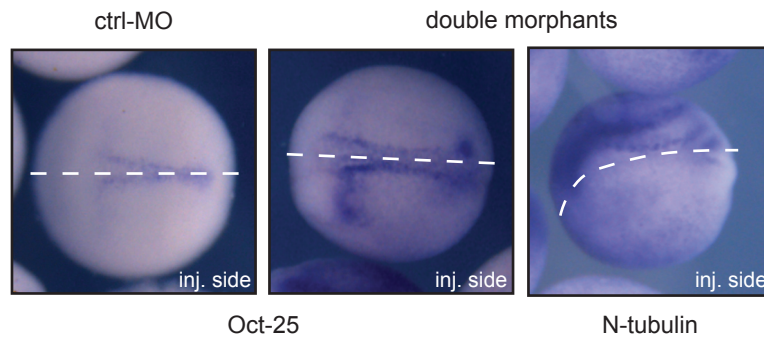


C

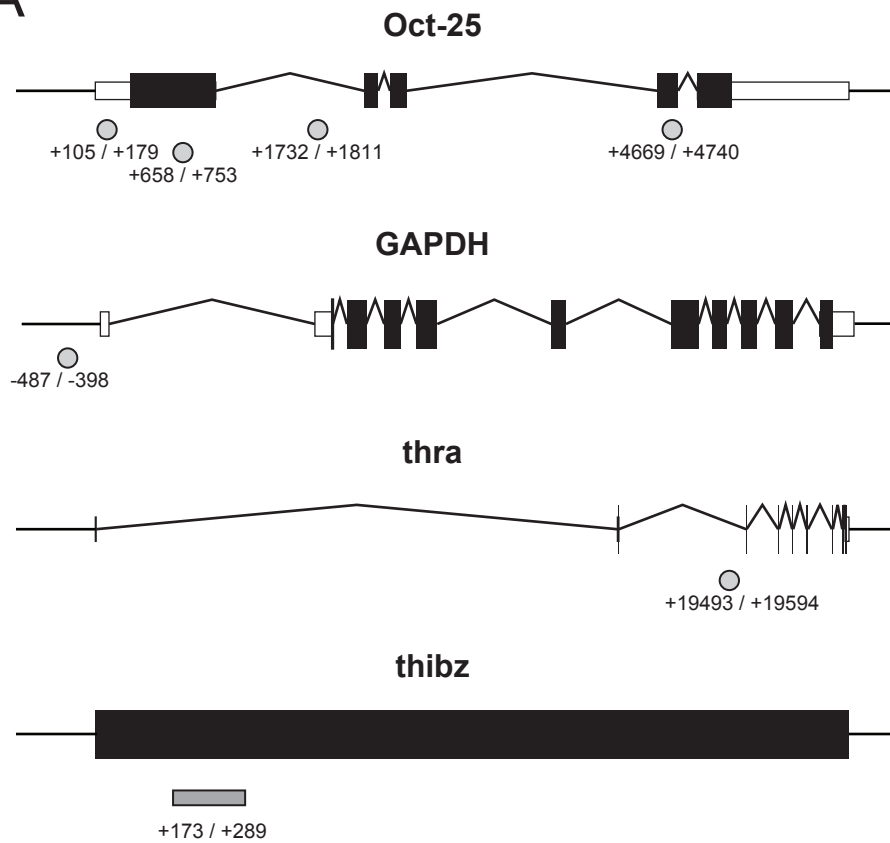


D

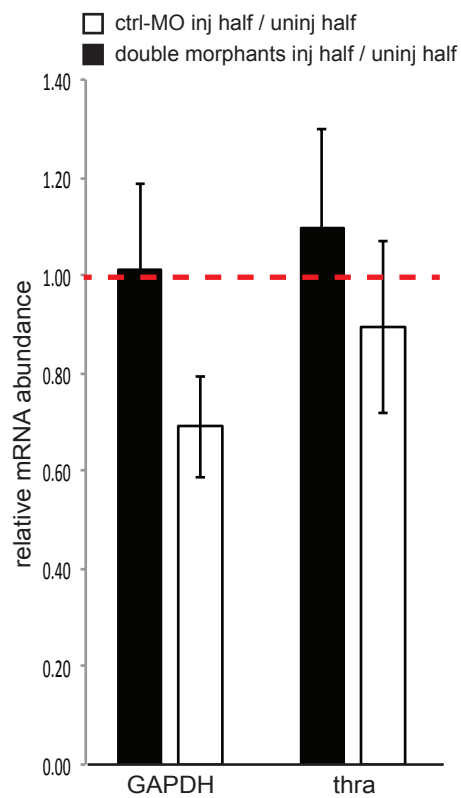
Top 10 Upregulated Genes			
Gene Name	Gene Symbol	logFC	Fold Change
Glucokinase (hexokinase 4)	gck	2.68077	6.41197
Rab interacting lysosomal protein	rilp	2.53108	5.78006
Tripartite motif containing 7	trim7	2.41965	5.35040
MGC81526 protein	MGC81526	2.40205	5.28555
Fat storage-inducing transmembrane protein 2	fitm2	2.23256	4.69967
Teratocarcinoma-derived growth factor 1	tdgf1	2.21938	4.65692
CDC42 small effector protein 2-c	cdc42se2-c	2.12568	4.36408
cAMP responsive element binding protein 1	creb1	2.09589	4.27490
POU class V protein Oct-25	pou5f1.1	1.97463	3.93028
Serine/threonine kinase 35	stk35	1.96083	3.89285

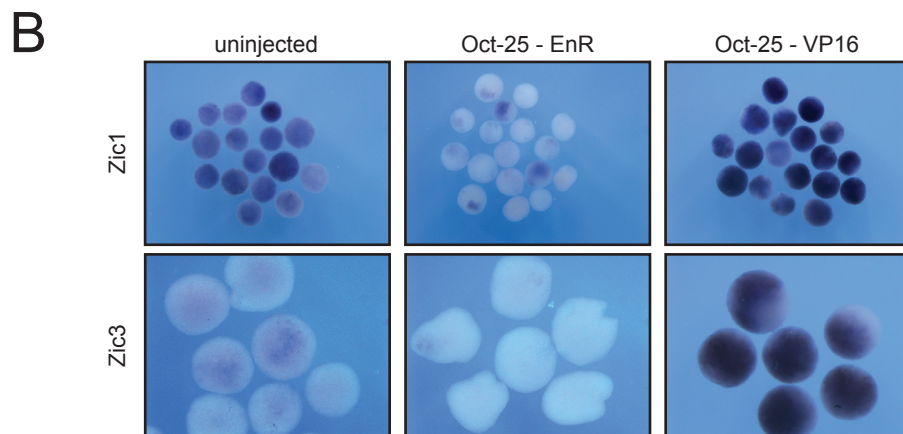
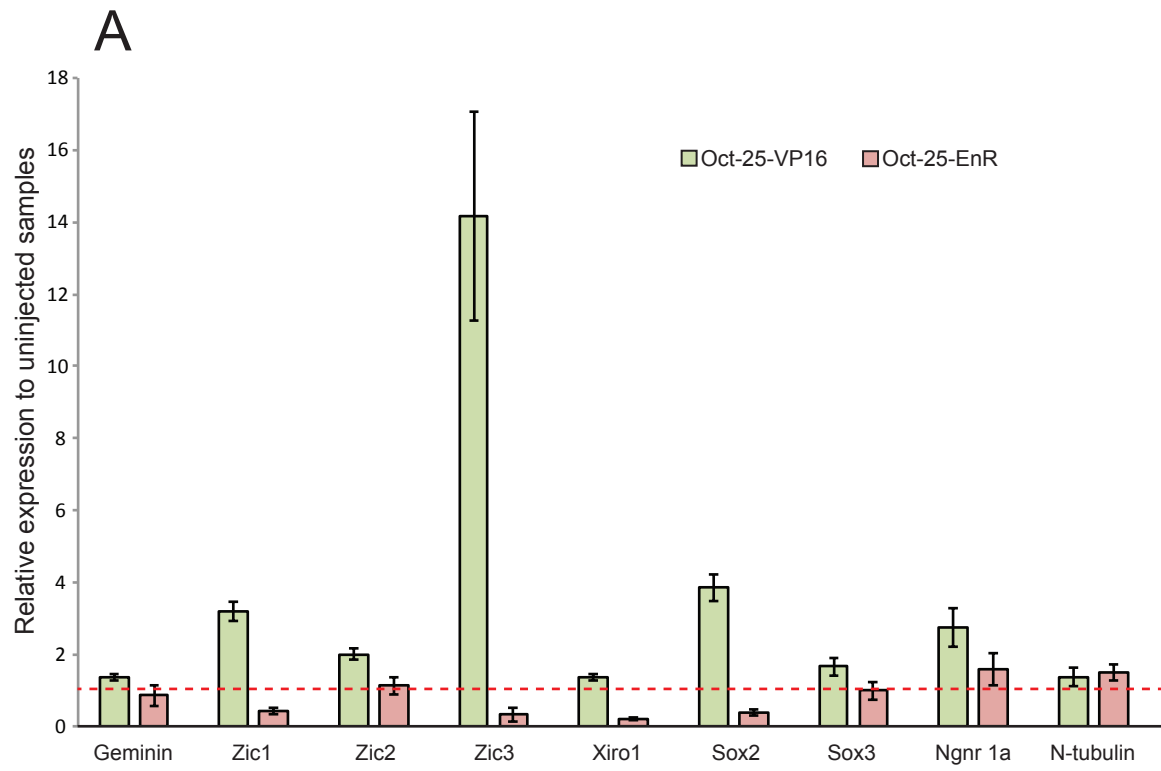


A

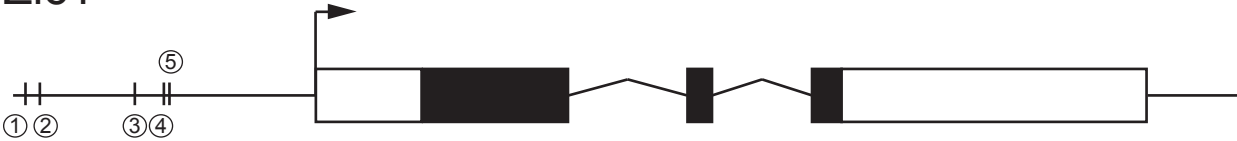


B





Zic1



Zic1				
start	end	sequence	weight	position
-1918	-1909	AATGGAAAAA	5.2	①
-1823	-1814	AATGCAACAA	6.8	②
-1195	-1186	TATGCAAAAT	6.6	③
-1003	-994	AATGCAATGT	6.7	④
-966	-957	CATGGAAAAT	5.3	⑤

Zic3

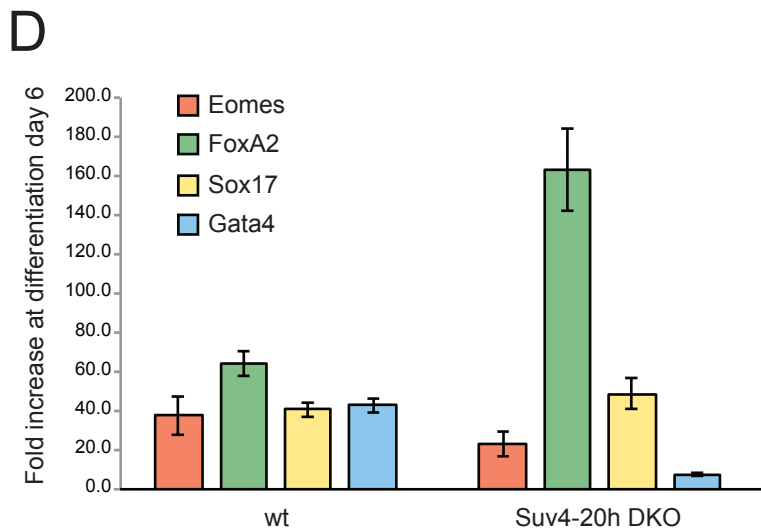
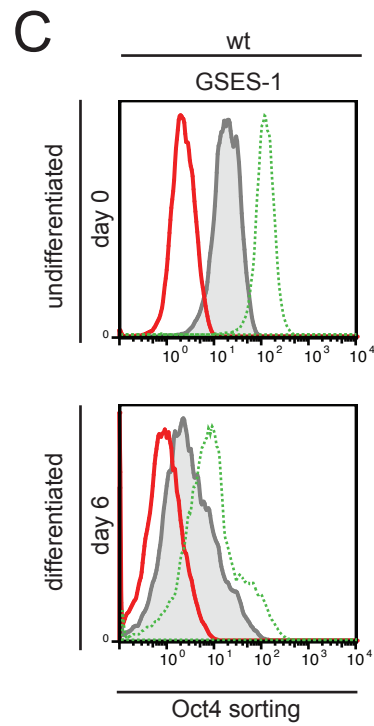
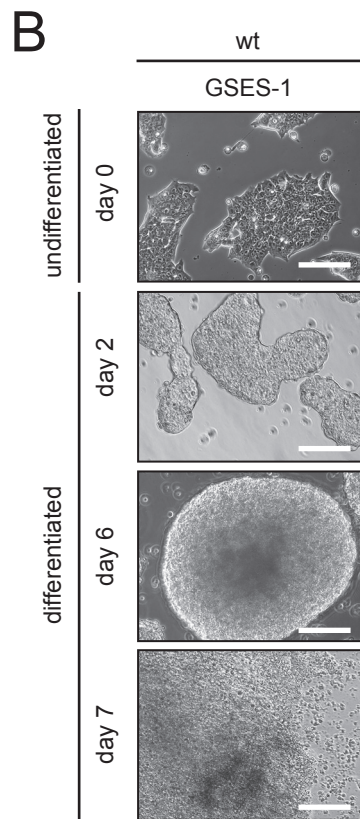
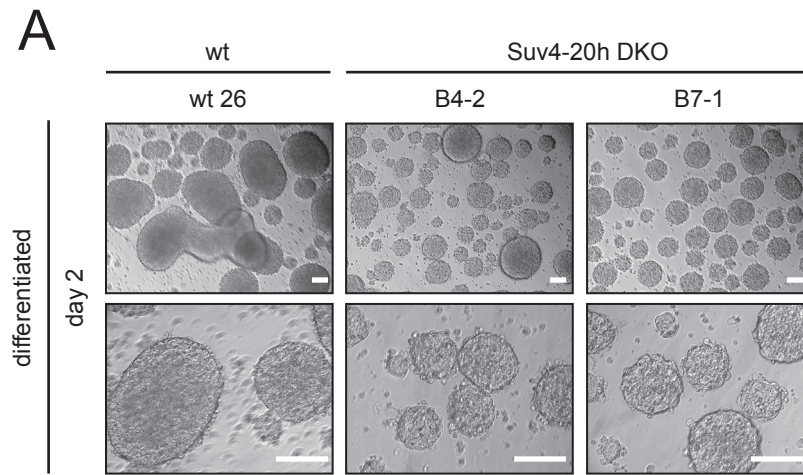


Zic3				
start	end	sequence	weight	position
-1816	-1807	AATGCAATAC	6.7	①
-1762	-1753	GATGCAAAGC	7.2	②
-1704	-1695	AATGCAAGCC	7.5	③
-1046	-1037	TATGCAAGTC	6.7	④
-928	-919	AATGCAAGAT	6.9	⑤
-252	-243	GATGCAAAAT	6.9	⑥

Sox2



Sox2				
start	end	sequence	weight	position
-1803	-1794	CATGAAATAA	5	①
-1510	-1501	GATGTAAGAG	5.1	②



OLIGONUCLEOTIDES SEQUENCES

RT-PCR - qRT-PCR (Xenopus)

Geminin for	5'-tgaagtggctgttgatccag-3'
Geminin rev	5'-tcttcgttcctctgcaacct-3'
H4 for	5'-gaccgcggtcacctacacc-3'
H4 rev	5'-ctggcgcttcagaaacataca-3'
MyoD for	5'-aggaaggccgccactatga-3'
MyoD rev	5'-gttgcgaggatctccactt-3'
Ngnr 1a for	5'-acctgcactctgcgcttgat-3'
Ngnr 1a rev	5'-gcgcaaggtctcatcttg-3'
Nrp1 for	5'-gccatgctgcaaaactctt-3'
Nrp1 rev	5'-cccaccttatagccctccat-3'
N-tubulin for	5'-tgctgatctacgcaaactgg-3'
N-tubulin rev	5'-ctgtcagggctcggtattgt-3'
Oct-25 for	5'-caggttcagggtgcag-3'
Oct-25 rev	5'-gtccttgagggtcaggaaag-3'
Oct-91 for	5'-ggacaacagtcgctgtagca-3'
Oct-91 rev	5'-cactgctcagcccatcacta-3'
ODC for	5'-acaaagaaacccaaaccaga-3'
ODC rev	5'-caaacaacatccagtctcaa-3'
Sox2 for	5'-tgcgtccaacaaccagaata-3'
Sox2 rev	5'-agttgtgcatctgggggttc-3'
Sox3 for	5'-atgaacggctggactaatgg-3'
Sox3 rev	5'-tacctgtgctggatctgtg-3'
Sox11 for	5'-cgagaaaatccccttcatca-3'
Sox11 rev	5'-aggatccactttgggcttttc-3'
Sox17 alpha for	5'-tactgcaactacccagtc-3'
Sox17 alpha rev	5'-agagcccgtccttcaata-3'
Xiro 1 for	5'-ccataaccaccaccaccttc-3'
Xiro 1 rev	5'-tgtctgagtgtgggactg-3'
XK81 for	5'-ccgttggtgtgaacaagtg-3'
XK81 rev	5'-gcagctcaattccaagctc-3'
xlSuv4-20h1 for	5'-gttgcatgaagtgggttg-3'
xlSuv4-20h1 rev	5'-gcagacaatcggttccatt-3'
xlSuv4-20h2 for	5'-ccgatgttcttcagaga-3'
xlSuv4-20h2 rev	5'-ccaccaggagtcaatctttc-3'
Zic1 for	5'-acagatgaggctgggcttc-3'
Zic1 rev	5'-cagttggctggaggcataat-3'
Zic2 for	5'-tcggtaggacggagcaatac-3'
Zic2 rev	5'-ttcataggggagtactgggttg-3'
Zic3 for	5'-ggtggtgcagccttaactc-3'

Zic3 rev 5'-tggcacaaaagtccatgttga-3'

ChIP-qPCR

GAPDH for	5'-ctgtgctactgggtctttcc-3'
GAPDH rev	5'-taagcacaggcagcccttac-3'
Oct-25 5'-UTR for	5'-ctccgacttattgggtgga-3'
Oct-25 5'-UTR rev	5'-tctaacctggatgggaggtg-3'
Oct-25 exon 1 for	5'-agagtccccagaacccaaat-3'
Oct-25 exon 1 rev	5'-aagggtaccagtccatgtg-3'
Oct-25 intron 1 for	5'-aaagctaccggctgattgg-3'
Oct-25 intron 1 rev	5'-agcgtgcaggattaggtcat-3'
Oct-25 exon 4 for	5'-aggggacgctggaaagttac-3'
Oct-25 exon 4 rev	5'-ccttggctattgcaccatc-3'
MSAT 3 for	5'-ccaccgtttgtcgtagacc-3'
MSAT 3 rev	5'-tgctggggcaattaactg-3'
Thibz for	5'-gctgtcggaaactctcactcc-3'
Thibz rev	5'-gcgtctctgtccagtagc-3'
Thr alpha for	5'-atttgcttcatgccttgct-3'
Thr alpha rev	5'-tatgaaacggagcgacacaa-3'

mutagenesis PCR

mSuv4-20h1 Y299A for	5'-cctggagaagaaatttctgttacgcaggagatggctttttggagaaa-3'
mSuv4-20h1 Y299A rev	5'-tttctccaaaaagccatctcctgcgtaacaagaaatttcttccagg-3'
mSuv4-20h2 Y217A for	5'-ggatgaagtgacttgcttcgagggtgagggcttctcgg-3'
mSuv4-20h2 Y217A rev	5'-ccgaagaagccctcacctgcgaagcaagtcacttcaccc-3'
mSuv4-20h1 N264A for	5'-ggctcggctcgtgctgatttatagcccatgattgcagacctaactg-3'
mSuv4-20h1 N264A rev	5'-cagttagggtcgtcaatcatgggctataaatgcagcaggaccgagcc-3'
mSuv4-20h2 N182A for	5'-ggcccagctgccttcacgcccagctgcaaacc-3'
mSuv4-20h2 N182A rev	5'-gggttgcagtcagggcgatgaaggcagctgggcc-3'

qRT-PCR (ES cells)

Actin for	5'-ggcatcactattggcaacg-3'
Actin rev	5'-tccataccaagaaggaagg-3'
Eomesodermin for	5'-atcgaccataacccctcgcc-3'
Eomesodermin rev	5'-cgtaccgacctccagggacaac-3'
FoxA2 for	5'-gtgaagatggaaggacagagc-3'
FoxA2 rev	5'-gccgcggacatgctcatgta-3'
GAPDH for	5'-tcaagaagggtggtgaagcag-3'
GAPDH rev	5'-gttgaagtgcaggagacaa-3'
Gata4 for	5'-agggtgagcctgtatgtaatgc-3'
Gata4 rev	5'-attcaggttctgggctcc-3'

Sox17 for	5'-acgcaagcggttggcacag-3'
-----------	---------------------------

Sox17 rev	5'-cgaagggccgcttctctgc-3'
-----------	---------------------------

SUPPLEMENTARY INFORMATIONS ON EXPERIMENTAL PROCEDURES

Statistical analysis

The statistical analysis was performed using two-tailed, Paired Student's *t*-test, unless differently specified.

Myc-tagged fusion protein extraction from embryos

25 embryos per condition were lysed in 100µl of 100mM NaCl, 10mM Tris pH 7.5 buffer supplemented with 1mM NaF, 20mM beta-glycerol, 0.1mM Sodium Vanadate, 10mM Na Butyrate, 0,5% NP-40 and EDTA-free protease inhibitor cocktail tablets (Roche). Embryos were centrifuged 15min at 14,000g at 4 °C; the supernatant was collected and 2X Loading buffer (Roti-Load1; carlroth.de) was added. Samples were subsequently analysed by western blot.

qRT-PCR samples preparation

Two-cell stage embryos were injected with ctrl-MO (80ng) or xSuv4-20h1 and h2 morpholinos (40ng each) mixed with Alexa 488 in only one blastomere only. At neurula stage (NF15), injected embryos were cut along the midline into pools of injected and uninjected halves based on alexa fluorescence. As control, embryos unilaterally injected only with Alexa 488-Dextran were processed in parallel. Six halves each from corresponding embryos were pooled into injected and uninjected sample pairs and used for RNA extraction. RNA samples from two independent experiments were subjected to microarray analysis, while RNA samples from four independent replicates were used to perform qRT-PCR analysis. For qRT-PCR profiles of ES cell lines, 10⁶ cells at day 0 and embryoid bodies at day 6 of differentiation were harvested and stored at -80°C. Total cellular RNA was extracted from independent experiments. For both frog and mouse samples, cDNA synthesis was performed using DyNAmo cDNA synthesis kit (Finnzymes), following the manufacturers' protocol. For microarray analysis, cDNA probes were prepared according to standard Affimetrix protocol. For each experiment a control RNA aliquot was processed without reverse transcriptase (–RT sample). Real time PCR was performed using the Power SYBR Green PCR Master mix and run in LightCycler[®] 480 System (Roche). Primer sequences are listed in Supplement Table S1. C(t) values for each sample were normalized to histone H4 as reference gene. The fold change between samples was then calculated by normalizing ctrl-MO injected embryos or xSuv4-20h double-morphant embryos to the uninjected samples by the

$\Delta\Delta C(t)$ method. Finally the ratio between injected and uninjected side within each sample was calculated to estimate the up- or downregulation of a gene's mRNA level. For qRT-PCR profiles of ES cell lines, $C(t)$ values for each sample were normalized to the two reference genes, GAPDH and Actin. $\Delta\Delta C(t)$ method was applied to calculate fold change difference between samples.

Vibratome sections of Oct-25 stained embryos

Embryos were rinsed in gelatine/albumin mixture (2.2g of gelatine dissolved in 500ml 1X PBS subsequently supplemented with 135g of albumin (Roth) and 90g of Sucrose). 1/20 vol of glutaraldehyde were added to 2ml of albumin/gelatine mixture. The solution was quickly vortexed and poured in a small plastic tray to create a bottom layer. Embryos were placed and properly oriented on the solidified layer. A second layer of albumin/gelatine mixture plus glutaraldehyde was prepared and poured on the embryos. The mixture was hardened at least for 30min. The gelatinized block with embedded embryos was cut out under a dissecting microscope and glued onto a metal support. 30-50 μ m sections were created using a Vibratome 1000 (Technical Products International, INC.). Sections were transferred on slides, slightly dried, covered with X-TRA Kit mounting medium (Mediate) and analysed with Leica M205FA Fluorescence Stereomicroscope.

Immunostaining

Immunostaining was performed according to Sive et al. (2000). Chromogenic reactions with BCIP/NBT (biomol) were stopped by rinsing embryos in PBS. Embryos were refixed in MEMFA and bleached in 1% H_2O_2 , 5% Formamid, 0.5X SSC on a light box for at least 4 hours. For immunostaining of paraffin embedded samples, embryos were fixed in MEMFA for one hour at room temperature and then transferred in ice-cold Dent's Fixative o/n at -20°C. Prior embedding embryos were rehydrated for 30min in 100mM NaCl, 100mM Tris/HCl pH 7.4. After dehydration with increasing ethanol concentrations, embryos were incubated for two hours in Xylene. Subsequently embryos were soaked in paraffin at 55°C twice for two hours, followed by proper orientation in moulds and paraffin hardening on cooling plates. Embryos were sectioned into slices of 10 μ m, which were dried on glass slides for 2 hours at 37°C. Paraffin was removed washing the samples twice with X-tra Solv (Mediate), then with decreasing ethanol concentration and finally with 1X PBS. Heat-induced epitope retrieval was performed incubating the slides in citrate buffers solution for 1 hour at 90°C followed by cooling down to room temperature. Endogenous peroxidase

inactivation was achieved by 10min incubation with 3% peroxidase inactivating solution (35% hydrogen peroxidase, Roth, 1/10 Methanol in PBS). Unspecific antibody binding sites were blocked by incubation for 1 hour with 2% biotin-free albumin (Roth) in PBS. Primary antibodies, diluted in blocking-solution, were incubated o/n at 4°C. Secondary antibody incubation was preceded by washes in PBST (1X PBS + 0.1% Tween20); subsequently, slices were incubated 1 hour at room temperature with biotinylated anti-Rabbit secondary antibody. After several washes in PBST, slices were incubated for 1 hour in the dark at room temperature in High Sensitivity Streptavidin-HRP solution (Thermo Scientific), diluted 1:500 in blocking solution. Staining was stopped after about 10min at room temperature in DAB substrate chromogen solution by washing the samples in double distilled water. For counterstaining haemalaun (Roth) was used (6 min at room temperature in 1:3 haemalaun-solution); slices were then blued with 10min under running tap water. Increasing ethanol concentrations and X-tra Solv were used for dehydration. Finally slides were embedded using X-TRA Kit mounting medium (Mediate) and analyzed with Leica DM microscope.

Immunofluorescence microscopy of MEF cells

Wt or Suv4-20 DKO MEFs were fixed with 3.7% (w/v) paraformaldehyde in PBS for 10min at room temperature, PBS washed and permeabilized in 0.1% (w/v) NaCitrate containing 0.1% (v/v) Triton X-100 for 5min. After washing in PBS-T (PBS, 0.1% [v/v] Tween-20), cells were blocked for 30min in blocking buffer (PBS, 2.5% [w/v] BSA, 0.1% [v/v] Tween-20) at room temperature. The cells were incubated with primary antibody (H4K20me3, diluted in blocking buffer; [24]) overnight at 4°C. After washing with PBS-T, the secondary antibody (Cy3-conjugated anti-rabbit; Jackson ImmunoResearch, diluted in PBS-T) was added for 1 hour at room temperature. After several washes in PBS-T, the cells were embedded in Vectashield containing DAPI (Vector Laboratories) and stored at 4°C for further analysis.

Chromatin Immunoprecipitation experiments

Aliquots of 50 *Xenopus tropicalis* injected and uninjected embryos were fixed at NF 14-15 in 5ml 1% formaldehyde in PBS for 5min at 20°C on a rolling wheel. Crosslinking was stopped by a 10min wash with 0.125M glycine/PBS, followed by three washes in PBS. Fixed embryos were transferred in 1.5ml eppendorf tubes, frozen in liquid nitrogen and stocked at -80°C. At experimental day1, embryos were thawed for 15min on ice.

Two 15ml conical tubes of blocked protein-G and -A (Protein-A and -G Sepharose 4, Fast Flow, GE Healthcare) beads were prepared by incubating the proper amount of beads (plus an extra 50 μ l) with 15ml of 5% BSA in PBS. The tubes were incubated at 4 °C while mixing for at least 1 hour.

Two 50-embryos aliquots (100 embryos/condition) were used in each experiment. 600 μ l of 4°C RIPA [49] buffer was added to each 50 embryos aliquot. Samples were homogenized with a pellet pestle by gently disrupting the embryos until no large embryo fragments are visible. Embryos were incubated on ice at least 10min and subsequently centrifuged at 14,000rpm for 10min at 4°C. The supernatant was discarded and the wall of the tubes was wiped with a kimwipe to remove lipid residue. 650 μ l of 4°C RIPA buffer was added to each sample; the pellet was re-homogenized vigorously. Samples were subsequently sonicated using the Bioruptor (Diagenode) for 25 cycles each composed by 30sec pulse and 30sec rest. Samples were centrifuged at 14,000rpm for 10min at 4°C. 600 μ l sheared chromatin from the two 50-embryos aliquots per sample were pooled together and transferred into a pre-chilled, clean 1.5 microcentrifuge tube. Input samples were then prepared: combining 5 μ l of sheared chromatin plus 195 μ l TES [49]. Input sample was frozen at -80°C and processed once the immunoprecipitations were completed.

One of the two 15ml conical tubes containing the blocked protein-G and -A beads was centrifuged at 1000rpm for 5min at 4°C. Excess of 5% BSA/PBS was removed and the beads were gently resuspended by pipetting. Pre-clearing step was achieved by dispensing 50 μ l blocked beads to each sample of sheared chromatin and incubating each sample at 4°C with mixing for 1-1.5 hour. Samples were subsequently centrifuged at 1000rpm for 1min at 4°C. Each 1.2ml sheared chromatin sample was separated into two samples by transferring 580 μ l of pre-cleared, sheared chromatin in two new 1.5ml pre-chilled, clean microcentrifuge tubes. Each new tube was filled with RIPA buffer and the immunoprecipitation was achieved by adding the appropriate amount of antibody to only one of the two tubes, keeping the second one as negative control. Samples were incubated overnight at 4°C with mixing.

At experimental day2 the second 15ml conical tube containing the blocked protein-G and -A beads was centrifuged at 1000rpm for 5min at 4°C. Excess 5% BSA/PBS was removed, the beads were gently resuspended by pipetting. 50 μ l blocked beads was added to each sample. Samples were incubated at 4°C with mixing for 1.5 hour and afterwards centrifuged at 100rpm for 1min at 4°C. Beads were subsequently washed: each wash consisted of a 1-minute spin at 1000rpm in a 4°C centrifuge to

pellet the immunocomplexes, removal of supernatant with a 20-gauge needle, addition of 1ml wash buffer and incubation at 4°C with mixing for 5min. Samples were washed for a total of 8 times, using 2 washes each with buffers I through IV. Following the washes, the supernatant was aspirated with a 26-gauge needle inserted into the beads to completely remove any residual wash buffer. 200µl TES buffer was added to the beads. Elution was achieved by incubating the samples at 65°C for 1 hour in a table shaker (1000rpm). During this time the frozen input samples were thawed and vortexed to resuspend any precipitated SDS. All the different samples (input, IP and negative control) were processed in the same manner for the rest of the procedure.

After elution samples were centrifuged at 14,000rpm at RT for 1min. 200µl of the eluted supernatant was transferred to a new 1.5ml microcentrifuge tube. RNase treatment was achieved by adding 2µl of 10 mg/ml stock RNase A (Qiagen) to each sample and by incubating the samples for 45min at 37°C. Subsequently, 12µl of Proteinase K/Glycogen solution was added to each sample. Samples were incubated at 68°C for 4 hours while shaking (1300rpm) to reverse crosslinks and digest proteins. DNA was purified on column using the QIAquick® PCR purification kit (Qiagen) following the manufacturer's protocol. DNA was eluted in 33µl of EB buffer. qPCR was performed using the Light Cycler 480 System (Roche). Data were analysed using the Light Cycler 480 Software Release 1.5.0 SP1 (Roche).

2.4 H3K56me3 is a novel, conserved heterochromatic mark that largely but not completely overlaps with H3K9me3 in both regulation and localization.

Article: Jack et al. 2013, PLOS ONE

H3K56me3 Is a Novel, Conserved Heterochromatic Mark That Largely but Not Completely Overlaps with H3K9me3 in Both Regulation and Localization

Antonia P. M. Jack¹, Silva Bussemer¹, Matthias Hahn¹, Sebastian Pünzeler¹, Martha Snyder², Michael Wells², Gyorgyi Csankovszki², Irina Solovei³, Gunnar Schotta¹, Sandra B. Hake^{1*}

1 Center for Integrated Protein Science Munich (CIPSM) at the Adolf-Butenandt-Institute, Department of Molecular Biology, Ludwig-Maximilians-University Munich, Munich, Germany, **2** Department of MCDB, University of Michigan, Ann Arbor, Michigan, United States of America, **3** LMU Biozentrum, Department of Biology II, Ludwig-Maximilians-University Munich, Planegg-Martinsried, Germany

Abstract

Histone lysine (K) methylation has been shown to play a fundamental role in modulating chromatin architecture and regulation of gene expression. Here we report on the identification of histone H3K56, located at the pivotal, nucleosome DNA entry/exit point, as a novel methylation site that is evolutionary conserved. We identify trimethylation of H3K56 (H3K56me3) as a modification that is present during all cell cycle phases, with the exception of S-phase, where it is underrepresented on chromatin. H3K56me3 is a novel heterochromatin mark, since it is enriched at pericentromeres but not telomeres and is thereby similar, but not identical, to the localization of H3K9me3 and H4K20me3. Possibly due to H3 sequence similarities, Suv39h enzymes, responsible for trimethylation of H3K9, also affect methylation of H3K56. Similarly, we demonstrate that trimethylation of H3K56 is removed by members of the JMJD2 family of demethylases that also target H3K9me3. Furthermore, we identify and characterize mouse Mjmd2E and its human homolog HKDM4L as novel, functionally active enzymes that catalyze the removal of two methyl groups from trimethylated H3K9 and K56. H3K56me3 is also found in *C. elegans*, where it co-localizes with H3K9me3 in most, but not all, tissues. Taken together, our findings raise interesting questions regarding how methylation of H3K9 and H3K56 is regulated in different organisms and their functional roles in heterochromatin formation and/or maintenance.

Citation: Jack APM, Bussemer S, Hahn M, Pünzeler S, Snyder M, et al. (2013) H3K56me3 Is a Novel, Conserved Heterochromatic Mark That Largely but Not Completely Overlaps with H3K9me3 in Both Regulation and Localization. PLoS ONE 8(2): e51765. doi:10.1371/journal.pone.0051765

Editor: Janet F. Partridge, St Jude Children's Research Hospital, United States of America

Received: August 1, 2012; **Accepted:** November 7, 2012; **Published:** February 22, 2013

Copyright: © 2013 Jack et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This study was supported by grants from the DFG (HA 5437/3-1 and SFB TR5), as well as CIPSM to S.B.H. and from the National Institutes of Health (NIH) R01 GM079533 to G.C. I.S. was supported by DFG (SO1054/1) and BioImaging Network Munich and G.S. by SFB TR5, SPP1356 and BMBF (EpiSys). A.P.M.J. and S.P. are members of the IMPRS/LS program. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: sandra.hake@med.uni-muenchen.de

Introduction

Histones, the building blocks of chromatin, are subject to several posttranslational modifications including methylation, acetylation and phosphorylation that carry important functional information [1]. Over the last decades, it has become increasingly obvious that such chemical histone tags contribute to the regulation of DNA-related processes in a highly selective and specialized manner [2]. These posttranslational histone modifications (PTMs) either change nucleosome structure directly by affecting histone-DNA contacts or indirectly by recruiting PTM-binding proteins that act on the underlying chromatin structure, as has been proposed in the “histone code” hypothesis [3]. Although most marks are found on the flexible histone tail regions, some modifications have also been identified on core residues. One such core PTM, histone H3 lysine 56 acetylation (H3K56ac) [4], occurs in the α -N-helical region near the entry-exit sites of the DNA superhelix and is conserved from yeast to man [5]. It is most abundant during S phase [6,7] and has been shown to play a pivotal role in DNA damage response [6], chromatin integrity [8,9] and replication-coupled nucleosome assembly [10]. In a previous mass spectrometry-based study, we

were not only able to verify the existence of H3K56 acetylation in humans but were also able to identify low levels of mono- and trimethylation of lysine 56 on histone H3 (H3K56me1 and H3K56me3, respectively) [11]. Recently, it was demonstrated that monomethylation of H3K56 regulates DNA replication through interaction with the replication processivity factor PCNA and is catalyzed by the lysine methyltransferase (KMT) G9a (KMT1C) [12]. The involvement of H3K56me1 in such an important biological event led us to ask how trimethylation of this residue might be regulated and impact cellular processes. Despite the known *in vivo* existence of H3K56me3 [11], no further information concerning this novel histone H3 core modification has been established. We set out to learn more about its functional role by deciphering its chromatin localization and by identifying enzymes that set (“writer”) and erase (“eraser”) this mark.

Materials and Methods

Cell lines

Human HeLa Kyoto cells [13], and mouse C127 (ATCC CRL-1616) cell lines were grown in DMEM medium (PAA) supplemented

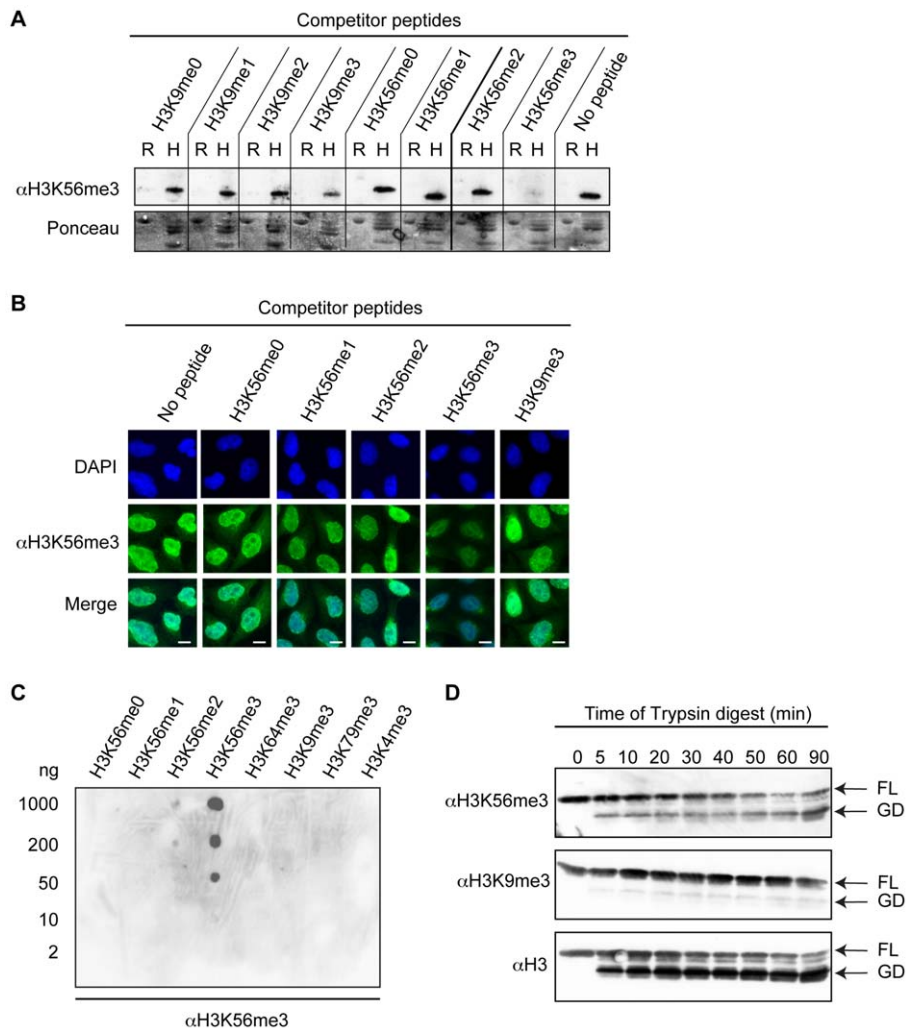


Figure 1. Determination of α H3K56me3 specificity and suitability in diverse applications. (A) Immunoblot peptide competition experiment. α H3K56me3 antibody was preincubated with competitor peptides before addition to immunoblots containing recombinant H3 protein (R) or acid extracted HeLa Kyoto histones (H) (top). Ponceau staining (bottom) serves as loading control. (B) IF microscopy peptide competition experiment. α H3K56me3 antibody (green) was preincubated with competitor peptides before addition to fixed HeLa Kyoto cells. DAPI (blue) stains DNA. Scale bar = 5 μ m. (C) Spot-blot with different concentrations (5–1000 ng) of H3 peptides to determine α H3K56me3-binding affinities. (D) Immunoblot of sequential tryptic digest of HeLa Kyoto-derived mononucleosomes using α H3K56me3 (top), α H3K9me3 (middle) and α H3 (bottom). FL = full-length histone H3, GD = N-terminally deleted globular domain of histone H3. doi:10.1371/journal.pone.0051765.g001

with 10% FCS (Sigma) and 1% penicillin/streptomycin at 37°C and 5% CO₂. Wild type, Suv39hDKO [14] and SUV4-20hDKO [15] mouse embryonic fibroblast (MEF) cell lines were grown in DMEM medium (PAA) supplemented with 18% FCS (Sigma), 1% penicillin/streptomycin, 1% non-essential amino acids (Invitrogen), 50 mM β -mercaptoethanol and 0.4% LIF at 37°C and 5% CO₂. Cells were transfected using FuGene HD (Roche Applied Science) according to the manufacturer's instructions.

Antibodies

Polyclonal rabbit antibody against H3K56me3 was developed by Pineda Antikörper-Service (Berlin, Germany) using a peptide with the following amino acid sequence for immunization and affinity purification: NH₂-CRRYQ-K(me₃)-STEL-CONH₂. Commercially available antibodies used in this study include: Primary antibodies: α H3 (C-terminus, Abcam), α H4 (Antikörper-online), α H3K4me2 (Abcam), α H3K4me3 (Abcam), α H3K9me1 (Millipore), α H3K9me2 (Active Motif), α H3K9me3 (Active Motif and

[16]; specificity tests are shown in Figure S1), α H3K27me2 (Millipore), α H3K27me3 (Millipore), α H3K36me1 (Millipore), α H3K36me2 (Active Motif), α H3K36me3 (Abcam), α H4K20me1 (Millipore), α H4K20me2 (Millipore), α H4K20me3 (Abcam), α H3K56me1 (Millipore), α H3K56me2 (Active Motif), α H3K56ac (Active Motif). Secondary antibodies: for immunoblots (Amersham), for IF microscopy (Dianova).

Peptide competition experiment

α H3K56me3 antibody in 1:1000 or 1:100 dilutions was preincubated with 2 μ g/ml of peptides (Table S1) before usage in either immunoblots or immunofluorescence (IF) microscopy, respectively. Peptides were N-terminally biotinylated and synthesized with higher than 80% purity by The Rockefeller University, GeneScript or the MPI for Biochemistry Munich. In case of immunoblots, acid extracted histones [17] and recombinant histone H3 [18] were used.

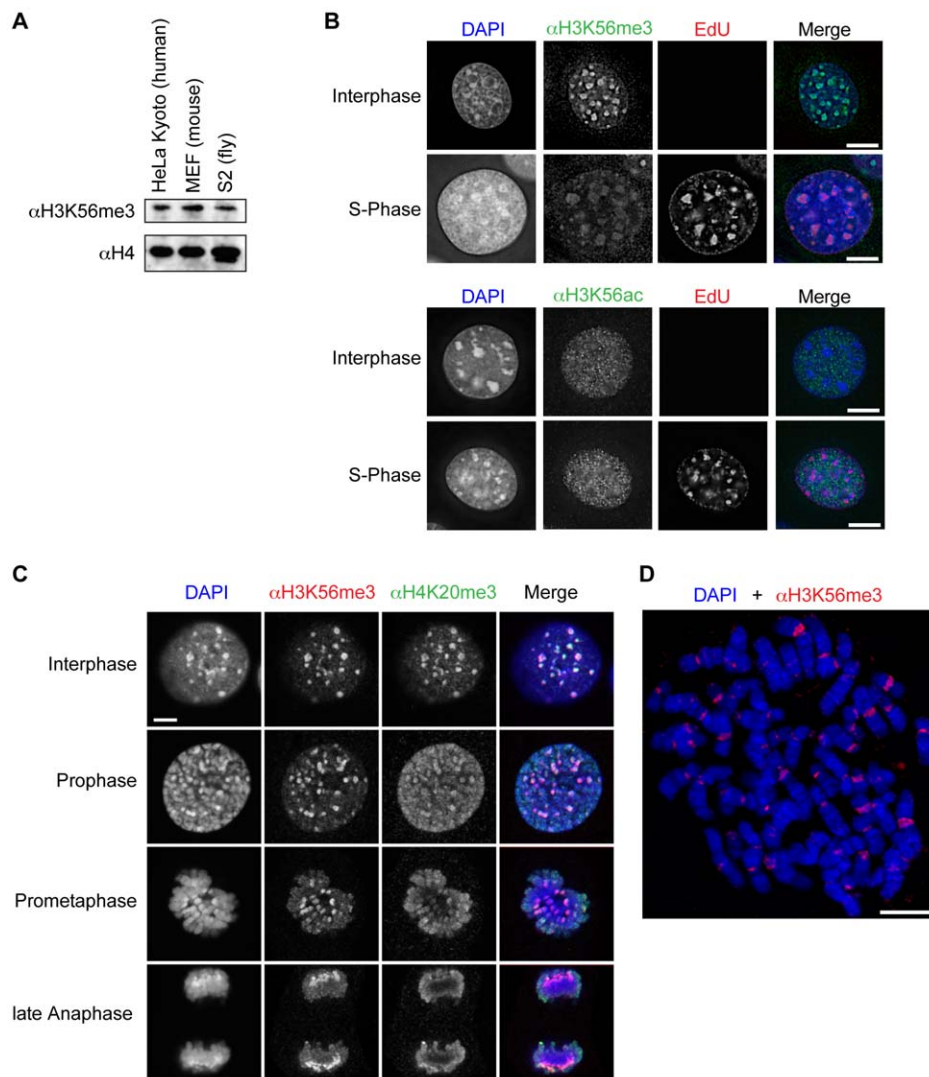


Figure 2. H3K56me3 is evolutionarily conserved, has a cell-cycle independent appearance and is part of pericentromeric heterochromatin. (A) Immunoblot with acid extracted histones from human (HeLa Kyoto), mouse (MEF) and fly (S2) cell lines using αH3K56me3 (top) and, as loading control, αH4 (bottom) antibodies. (B) IF analysis of H3K56me3 (top) and H3K56ac (bottom) appearance in G1/G2 and S-phase cells. C127 cells were pulse-labeled with EdU (red) to visualize replication foci and to identify cells in S-phase. Cells were co-stained with αH3K56me3 or αH3K56ac (green) and DAPI (DNA, blue). Scale bar = 5 μm. (C) IF microscopy of MEF cells in interphase and different stages of mitosis co-stained with αH3K56me3 (red), αH4K20me3 (red) and DAPI (DNA, blue). Scale bar = 5 μm. (D) IF of chromosome spread from nocodazole-arrested HeLa cells with αH3K56me3 (red) and DAPI (DNA, blue) staining. Scale bar = 5 μm.
doi:10.1371/journal.pone.0051765.g002

Tryptic digest of mononucleosomes

6×10^7 HeLa Kyoto cells were incubated in PBS, 0.3% Triton X-100 and Protease Inhibitor Cocktail (Roche, Germany) for 10 min at 4°C. Nuclei were pelleted, washed once in PBS, resuspended in EX100 buffer (10 mM Hepes pH 7.6, 100 mM NaCl, 1.5 mM $MgCl_2$, 0.5 mM EGTA, 10% (v/v) glycerol, 10 mM β-glycerol phosphate, 1 mM DTT, Protease Inhibitor Cocktail (Roche, Germany)) and $CaCl_2$ concentration adjusted to 2 mM. Resuspended nuclei were digested with 1.5 U MNase (Sigma) for 20 min at 26°C. The reaction was stopped by addition of EGTA to a final concentration of 10 mM followed by centrifugation for 10 min at 1000 rcf, 4°C. Mononucleosome containing supernatant was retained. NH_4HCO_3 was added at a final concentration of 50 mM or until a pH of 7–8 was reached. 1.6 μg Trypsin (Promega) was added and the reaction was incubated at 25°C. Samples were collected at different time

points and the reaction stopped by adding an equal volume of 1% trifluoroacetic acid. Fragments were size separated on a 15% SDS-PAGE probed with indicated antibodies.

Spot-blot

Peptide dilutions containing 2, 10, 50, 200 and 1000 ng in sterile water were spotted on nitrocellulose membrane and allowed to air-dry. The membrane was then blocked in PBS-Tween (0.1%) with milk powder (5%), followed by immunoblotting with αH3K56me3.

Immunofluorescence (IF) microscopy and cell cycle analysis

Mammalian cells. Preparation of mammalian cells and chromosome spreads for IF microscopy was done as previously reported [19]. Staining of S-phase cells was performed as

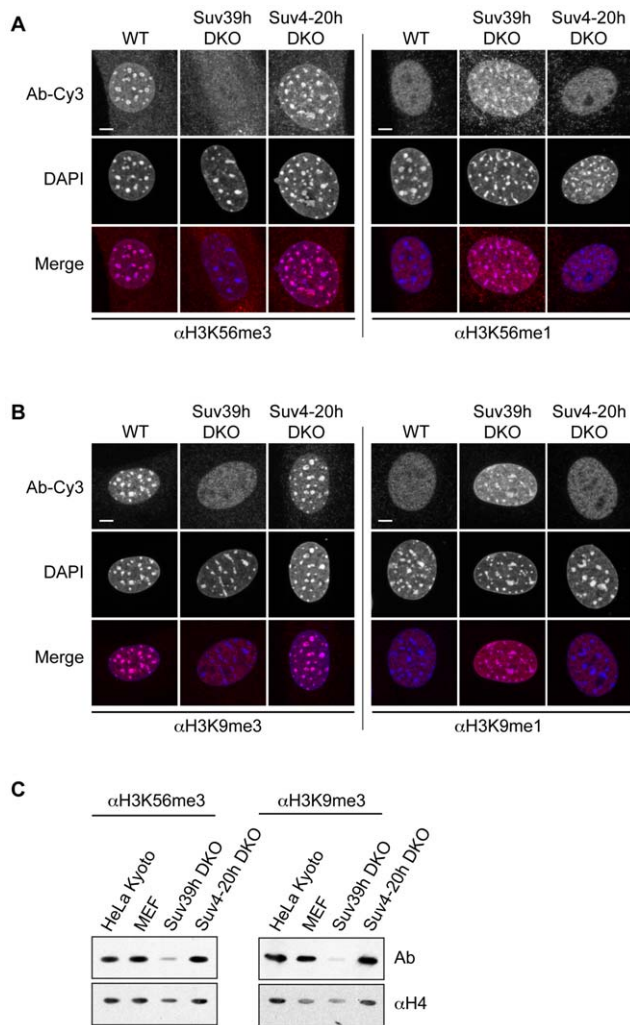


Figure 3. Loss of Suv39h enzymes affect H3K56me3. IF microscopy of wild type (WT), Suv39h double-null (Suv39h DKO) and Suv4-20h double-null (Suv4-20h DKO) MEF cells using various H3K56 (A) and H3K9 (B) methyl-specific antibodies (Ab-Cy3, red) and DAPI (DNA, blue). Scale bar = 5 μ m. (C) Immunoblots using acid extracted histones from HeLa Kyoto (positive control), wild type MEF, Suv39h DKO and Suv4-20h DKO cells. Blots were incubated with α H3K56me3 (left, top) or α H3K9me3 (right, top) antibodies, respectively. Blots shown at the bottom were incubated with α H4 to ensure equal loading. doi:10.1371/journal.pone.0051765.g003

described in [18]. Wide-field IF imaging of EdU-stained C127 cells was performed on a PersonalDV microscope system (Applied Precision) equipped with a 60 \times /1.42 PlanApo oil objective (Olympus), CoolSNAP ES2 interline CCD camera (Photometrics), Xenon illumination and appropriate filtersets. Iterative 3D deconvolution of image z-stacks was performed with the SoftWoRx 3.7 imaging software package (Applied Precision).

Confocal imaging of chromosome spreads was performed on a TCS SP5 II microscope system (Leica Microsystems, Wetzlar, Germany), equipped with a 63 \times /1.3 HCX PL APO glycerol immersion objective. Z-stacks were recorded and subsequently deconvolved with Huygens Essential Software (SVI, Hilversum, The Netherlands).

Image stacks of immunostained MEF cells were collected using a Leica TCS SP5 confocal microscope with Plan Apo Lambda Blue 63 \times /1.4 NA oil or 63 \times /1.3 glycerol immersion objective.

C. elegans. Methanol/acetone fixation for immunostaining was performed as follows. Adult hermaphrodites were dissected in 1 \times sperm salts with and frozen on dry-ice for 20–30 minutes. The slides were fixed in methanol followed by acetone, 2 minutes each wash, at -20°C . Slides were then washed once for ten minutes in PBST prior to incubation with primary antibody [1:200 or 1:100 (direct labeling) α H3K56me3, 1:1000 α H3K9me3 (Abcam ab8898)]. Remainder of staining protocol was conducted as described previously [20]. Microscopy and imaging were conducted as described previously [21].

Images were capture with a Hamamatsu Orca-Erga close-coupled-device (CCD) camera mounted on an Olympus BX61 motorized Z-drive microscope using a 60X APO oil immersion objective. These images are projections of optical sections with a Z spacing of 0.2 micrometers. Scale bars were added using ImageJ (available at <http://rsb.info.nih.gov/ij/>; developed by Wayne Rasband, National Institutes of Health, Bethesda, MD) and a template image created in Slidebook.

Quantitative PCR

qPCR was carried out as previously described [22] using Fast SYBR Green Master Mix (Applied Biosystems). Results were normalized to HPRT1 and GAPDH levels.

Cloning of GFP-jmjC constructs

pDONR entry clones of the Jmjd2 subgroup [23] were recombined into the target vector pEGFP-N1-GW using LR clonase II enzyme mix (Invitrogen) according to the manufacturer's protocol.

C. elegans RNAi

RNA interference by feeding was performed with the Ahringer laboratory RNAi library [24] in two generations as described previously [21].

Results

Development of a specific α H3K56me3 antibody

To gain insight into the biological function(s) of H3K56 trimethylation, we raised a polyclonal antibody against H3K56me3 (α H3K56me3) and determined its specificity in various assays. Since H3K56me1 has previously been reported to be catalyzed by the H3K9me1-specific KMT G9a, maybe due to a conserved lysine-serine-threonine (K/S/T) motif at the site of both residues [12], we put special emphasis on testing a potential cross-reactivity of this antibody with H3K9me3. First, we performed peptide competition experiments using peptides spanning diverse regions of histone H3 with or without different methylation states. Specific antibody recognition of H3K56me3 in immunoblotting (Figure 1A) and immunofluorescence (IF) microscopy (Figure 1B) was efficiently competed out only with H3K56me3-containing peptides, but not with peptides containing other methylated or unmethylated histone regions. Next, we determined the relative binding affinity of α H3K56me3 to its epitope by a peptide Spot-blot containing various concentrations of different histone peptides and observed that α H3K56me3 detected as low as 50 ng of H3K56me3 peptides (Figure 1C). Notably, α H3K56me3 does not recognize any other trimethylated peptides except H3K56me3. For further support of antibody specificity, we generated mononucleosomes from HeLa cells that were subsequently digested with different concentrations of Trypsin in order to generate histones lacking their flexible tail regions. In this way, we were able to determine if the antibody epitope resides in the H3 core region or N-terminal tail. In

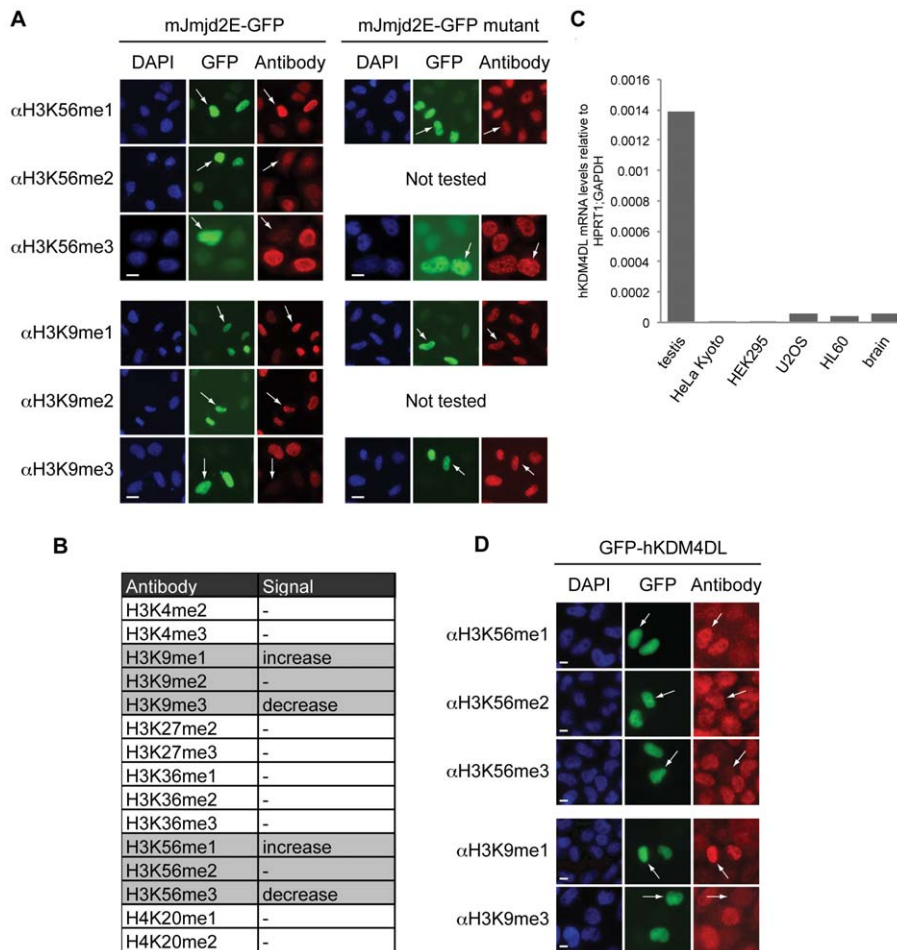


Figure 4. Jmjd2E demethylase affects H3K56me3. (A) IF microscopy of HeLa Kyoto cells transfected with mJmjd2E-GFP (green, left) or jmjcdomain mutated mJmjd2E-GFP (mutant, green, right) and stained with various H3K56 and H3K9 PTM-specific antibodies (red) and DAPI (DNA, blue). Arrows indicate transfected GFP-positive cells. Scale bar = 10 μ m. See also Figure S2A for IF results of cells transfected with other GFP-tagged mJMJD2 family members (mJmjd2a-d). (B) List of PTMs analyzed in IF after expression of mJmjd2E in HeLa Kyoto cells indicating changes in fluorescence intensities. See also Figure S2B for examples of IF results summarized in this table. (C) qPCR analysis with cDNAs from different human cell lines and tissues using primer pair specific for human Jmjd2E (hKDM4DL). Data were normalized to HPRT1 and GAPDH expression levels. (D) IF microscopy of HeLa Kyoto cells transfected with human Jmjd2E (hKDM4DL) (green) and stained with various H3K56 and H3K9 methyl-specific antibodies (red) and DAPI (DNA, blue). Arrows indicate transfected and GFP-positive cells. Scale bar = 10 μ m.
doi:10.1371/journal.pone.0051765.g004

immunoblots, α H3K56me3, but not the control α H3K9me3 antibody, recognized both full-length (FL) and the N-terminus deleted globular domain (GD) of histone H3 (Figure 1D), demonstrating that α H3K56me3 specifically binds to a modification in the core region of H3. In summary, these experiments provide compelling evidence that α H3K56me3 is highly specific for this particular modification and can be applied in diverse biochemical assays.

H3K56me3 is evolutionary conserved and localizes to pericentromeric heterochromatin outside of S-phase

Having demonstrated the high specificity of α H3K56me3, we first examined the evolutionary occurrence of this novel mark by isolating histones from cell lines of diverse origins. Immunoblotting revealed that H3K56me3 was present in human, mouse and fly (Figure 2A), suggesting that this modification is conserved within, at least, metazoans.

Given that H3K56ac is highly conserved and that methylation and acetylation of the same residue are mutually exclusive we

wanted to investigate if there were correlations between the appearance of one mark and disappearance of the other. While in yeast H3K56ac has been shown to be cell cycle dependent, showing a significant increase during S-phase, [6,9,25], its cell cycle distribution in mammals remains controversial [26–28], with a high possibility of its occurrence in all cell cycle phases [29]. Therefore, we analyzed cell cycle appearance and nuclear localization of both acetylation and methylation of H3K56 in mammalian cells. To distinguish S-phase from interphase, mouse C127 cells were pulse-labeled with the thymidine analog EdU, which was chemically coupled to a fluorescent dye using a “click-chemistry” approach [30]. Co-staining of EdU-labeled cells with α H3K56me3 revealed that, during interphase, H3K56me3 is found predominantly at DAPI-dense heterochromatic chromocenters and shows strongly diminished signal intensity in S-phase cells (Figure 2B top). Although, we observed a more or less equal appearance of H3K56ac signal in interphase and S-phase cells (Figure 2B, bottom), it is clearly distinct from the H3K56me3 signal. We also found H3K56me3 to be present throughout mitosis (Figure 2C), where it co-localizes with heterochromatin

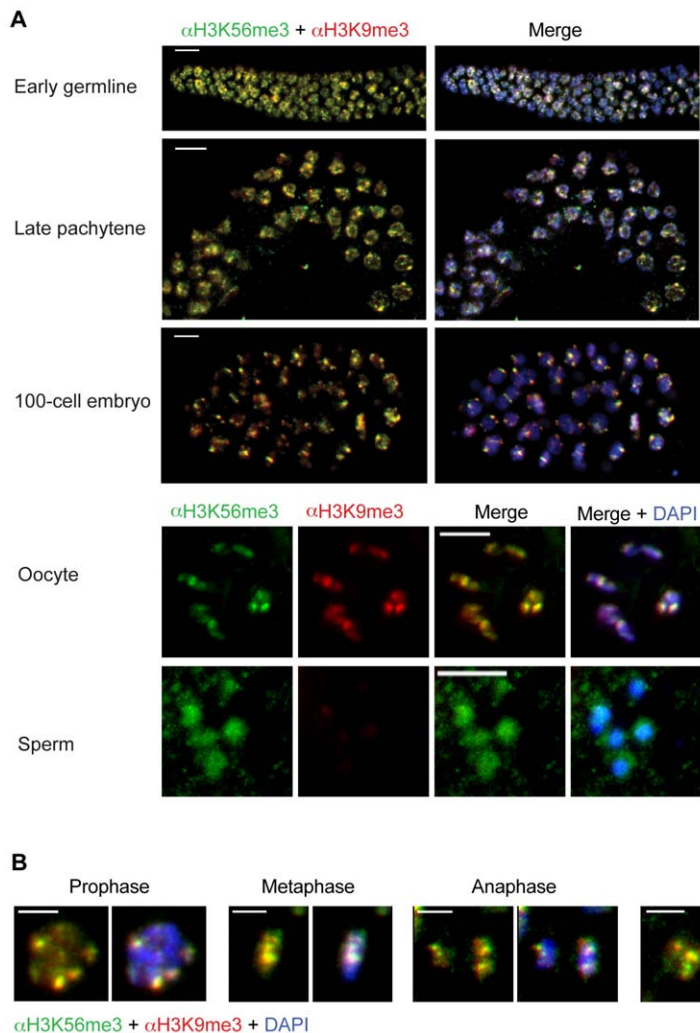


Figure 5. H3K56me3 is conserved in *Caenorhabditis elegans*. Shown are representative IF microscopy pictures from adult *C. elegans* hermaphrodite tissues. In all images H3K56me3 is shown in green, H3K9me3 in red, and DAPI (DNA) in blue. Scale bar = 5 μ m. A) H3K56me3 co-localizes with H3K9me3 in the early germline, late pachytene and in a 100-cell embryo (top picture). Interestingly, although H3K56me3 and H3K9me3 are both present in oocytes, only H3K56me3, but not H3K9me3, staining could be observed in sperm. (bottom, split channels) (B) H3K56me3 and H3K9me3 co-localize throughout all stages of mitosis.
doi:10.1371/journal.pone.0051765.g005

foci, in an even more precise manner than the constitutive heterochromatin marker H4K20me3 [31]. To determine H3K56me3 localization in greater detail, human metaphase chromosomes were analyzed in IF microscopy. In accordance with H3K56me3 presence at chromocenters in interphase and heterochromatin foci in mitotic cells, this modification was present in a non-random manner and found predominantly at pericentromeric heterochromatin regions that include major satellite repeats (Figure 2D). Interestingly, H3K56me3 is, in contrast to H3K9me3, rarely found at telomeres [32], suggesting that the functional roles of these two modifications in heterochromatic regions might be different.

Mammalian methyltransferase Suv39h affects trimethylation of H3K56

To assess the functional relevance of posttranslational histone modifications, it is important to know their responsible enzymes. Several lysine methyltransferases (KMTs) that catalyze the methylation of histone lysine residues have been identified

previously [33,34]. Possibly due to the fact that both regions surrounding H3K56 and H3K9 contain a conserved K/S/T motif, monomethylation of H3K56 has been shown to be catalyzed by the H3K9me1-specific KMT G9a [12]. Additionally, both H3K9me3 and H3K56me3 localize to similar, albeit not identical, nuclear domains suggesting that H3K9 and H3K56 might share the same KMTs responsible for their trimethylation. Therefore, we first tested the H3K9me3-specific KMTs Suv39h1/2 (KMT1A and B) [35] for their ability to affect the methylation status of H3K56. Interestingly, we observed a complete loss of both H3K56me3 and H3K9me3 signals at chromocenters in Suv39h double-null MEF cells (Suv39h DKO, [14]). Accompanied with this loss of trimethyl signals, we observed an increase of the respective monomethyl marks at chromocenters (Figure 3A and B). This dramatic change in PTM localization upon the simultaneous lack of Suv39h1/2 suggests that these enzyme are involved in catalyzing trimethylation of both H3K9 and H3K56, the latter in either a direct or indirect manner. Since H3K56me3 showed a somewhat similar nuclear appearance as H4K20me3

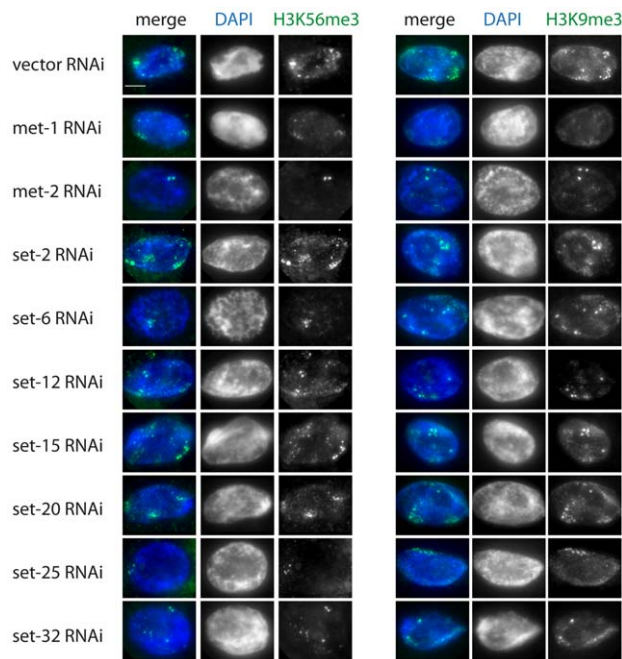


Figure 6. *C. elegans* RNAi screen to identify H3K56me3-specific KMTs. Shown are representative IF images from adult *C. elegans* hermaphrodite somatic intestinal nuclei following RNAi treatment. H3K56me3 (left) or H3K9me3 (right) staining is shown in green and DAPI (DNA) is shown in blue. CAPG-1 co-staining was used as a staining control (data not shown). Results show that *met-1* and *met-2* depletion severely affect both H3K56me3 and H3K9me3, while reduction of additional KMTs (*set-6*, *set-25* and *set-32*) has a stronger effect on H3K56me3 levels compared to H3K9me3. Scale bar = 5 μ m. doi:10.1371/journal.pone.0051765.g006

(Figure 2C), we wondered whether Suv4-20h1/h2 enzymes, responsible for methylating lysine 20 on histone H4 [15,31], might also target H3K56. Suv4-20h double null MEF cells (Suv4-20h DKO, [15]) showed no difference in abundance or localization of H3K56 methylation when compared to wild type cells (Figure 3A and B), demonstrating that these enzymes do not influence H3K56 methylation status. Similar results were also obtained with immunoblots, showing that the H3K56me3 signal is diminished in Suv39h DKO, but not Suv4-20 DKO acid extracted histones (Figure 3C).

Jmjd2E/KDM4DL is a novel lysine-demethylase specific for H3K9 and H3K56 trimethylation

Having shown that the same enzymes that methylate H3K9 also affect trimethylation of H3K56, we wondered whether the erasure of these modifications is catalyzed by identical lysine demethylases (KDMs) as well. Histone lysines are demethylated by two different classes of enzymes that are distinguished by their enzymatic active domains and methylation-state specificities [36]. We focused our attention on the Jumonji C-terminal domain (JmJC) family of KDMs, since they are able to remove all methyl-states, including trimethylation [37]. We therefore tested a panel of GFP-tagged members of the JMJD2 group-containing demethylases that are thought to partially work on H3K9me3 [23]. Over-expression of the respective mKDM in human cells was monitored by GFP signal in IF microscopy and effects on histone methylation were analyzed by co-staining with different histone PTM antibodies. This screen led to the identification of all members of the mJMJD2 family (mJmjd2A-E) able to affect H3K56me3 (Figure S2A and

Figure 4A). Since all members have previously been shown also to act on H3K9me3 [38], our results point once again towards a possible link between these two heterochromatic marks due, to a shared sequence motif (K/S/T). As one example, over-expression of mJmjd2D or mJmjd2E [39] strongly diminished H3K9me3, as well as H3K56me3 signals, in HeLa Kyoto cells (Figure 4A, left and Figure S2A). The loss of the respective trimethyl signal was accompanied with an increase in the monomethyl, but not dimethyl state, suggesting that these enzymes remove two methyl groups in total. Since over-expression of mJmjd2D-GFP caused severe cellular defects, its role on H3K56me3 was not further investigated and we focused subsequent analyses on mJmjd2E that acted solely on H3K9 and H3K56 trimethylation and not on other histone trimethylation marks (Figure 4B and Figure S2B). The observed changes in H3K9 and H3K56 methylation states upon mJmjd2E-GFP over-expression were dependent on the enzymatic active jmjC domain, since point mutations in that region completely abolished mJmjd2E's demethylase activity (Figure 4A, right). In the mouse, mJmjd2E is predicted to constitute a pseudogene and we therefore decided to analyze expression and function of the yet uncharacterized human homolog hKDM4DL. hKDM4DL mRNA is expressed predominantly in testis, with only residual levels present in U2OS (osteosarcoma) and HL60 (promyelocytic leukemia) cell lines and human brain tissue (Figure 4C). Over-expression of GFP-hKDM4DL in HeLa Kyoto cells showed identical results as seen for the mouse homolog, loss of H3K56 and H3K9 trimethylation with an accompanied gain of the respective monomethylation mark (Figure 4D). Taken together, we have identified the JMJD2 family to facilitate demethylation of H3K9 and H3K56 trimethyl states. Additionally, we showed that mJmjd2E, and its previously uncharacterized human homolog hKDM4DL, specifically remove two methyl groups from trimethylated H3K56 or H3K9 residues, depending on their catalytically active jmjC domain.

H3K56me3 is a novel chromatin mark in *C. elegans*

In order to learn more about H3K56me3 evolutionary conservation as a novel heterochromatic histone modification and its functions, we conducted IF microscopy analysis of wild type (WT) *C. elegans* hermaphrodite germlines and embryos (Figure 5A). H3K56me3 is present in both early germline and embryonic nuclei, as marked by DAPI morphology (Figure 5A, right). In almost all cells analyzed, we observed an H3K56me3 signal that strongly co-localized with H3K9me3 in most tissues (Figure 5A). Surprisingly, H3K56me3 staining was present in both types of germline cells, oocytes and sperm, whereas the H3K9me3 signal was restricted to oocytes only (Figure 5A, bottom). These data mirror previously obtained H3K9me3 results [40] and suggest that H3K56me3 might have an important H3K9me3-independent function in sperm development. Next, we wondered whether, similar to mammalian cells, H3K56 is trimethylated in cells during mitosis in *C. elegans*. Indeed, H3K56me3 is part of all mitotic stages and overlaps with H3K9me3 signals (Figure 5B), demonstrating the evolutionary high conservation of this novel mark.

Next, we sought to shed light on the enzymatic regulation of H3K56 trimethylation in *C. elegans* and performed an RNAi-based survey of known or predicted methyltransferases, including H3K9-specific enzymes [41]. The screen included RNAi targeting MET-2, a homolog of mammalian euchromatic H3K9 HMT SETDB1 [40–42], MET-1, a homolog of yeast Set2, an H3K36-specific methyltransferase, whose activity was reported to be required for normal levels of H3K9me3 [41], and SET-25, a distant homolog G9a, recently reported to deposit H3K9me3 in *C. elegans* embryos [42]. We also included RNAi against previously uncharacterized

SET domain containing proteins predicted to encode divergent H3K9-specific methyltransferases (set-6, -12, -15, -20, and -32). For control, we performed RNAi targeting an H3K4-specific methyltransferase SET-2, a homolog of SET1/MML. We conducted our screen in the intestine, where the large size of nuclei makes scoring easier. This screen identified several genes whose activity is required for normal levels of H3K56me3 and/or H3K9me3, some of which have been previously implicated in H3K9 methylation. H3K56me3 levels were severely reduced in *met-2* and *set-25* RNAi, consistent with the requirement for these genes for H3K9me3 levels in *C. elegans* embryos [42]. Interestingly, H3K9me3 levels were less affected in these conditions, indicating possible differences between the enzymes responsible for these marks and/or differences in antibody sensitivities. H3K56me3 levels were also reduced in *met-1* RNAi, and to a lesser extent in *set-6* and *set-32* RNAi. H3K9me3 levels were also reduced in *met-2* and *set-12* RNAi, possibly due to indirect effects (Figure 6). H3K9me3 levels were never reduced to background levels, perhaps due to partial redundancy between these enzymes. Knockdown of other known H3K9 methyltransferases or the H3K4 KMT *set-2* resulted in DAPI perturbations, but showed no effect on H3K56me3 staining (Figure 6).

In sum, H3K56me3, its relationship to H3K9me3, and its regulation by several H3K9 methyltransferases are conserved in *C. elegans*. However, some degree of divergence in the factors regulating H3K56me3 may have occurred in *C. elegans*.

Discussion

Our study establishes the existence of a novel pericentric heterochromatin mark, H3K56me3, in several metazoan species. This novel modification is present in all cell cycle phases, with the exception of S-phase, where it is underrepresented. Enzymes targeting H3K9 also act on H3K56, as the KMTs Suv39h1/2 are important for trimethylation of both residues and KDM JMJD2 family members remove these modifications. Mouse Jmjd2E and its so far uncharacterized human homolog hKDM4DL are involved in the process of demethylating H3K56me3 to a monomethylated status. In *C. elegans*, H3K56me3 is a conserved feature of mitotic chromosomes that primarily co-localizes with H3K9me3 and is regulated by some but not all H3K9 methyltransferases.

Of particular interest is our observation in mammalian cells that H3K56me3 is found in chromocenters containing pericentric heterochromatin, but only outside of S-phase. During that particular cell cycle phase, H3K56me3-specific IF microscopy signals are strongly diminished. Such an effect can be caused either by a replication-specific removal of the trimethylation mark or by occlusion of the epitope through adjacent modifications, such as phosphorylation of H3S57, or association with a binding protein. As H3K56 is targeted by the lysine acetyltransferases CBP or GCN5 [26,43] prior to being deposited onto DNA during replication [7,44,45], it is highly likely that newly synthesized H3 histones with K56ac replace “old” H3K56me3-containing ones. Given that H3K56me3 has been recently shown to prevent binding of PCNA that specifically associates with the monomethylation state [12], it is plausible that H3K56me3 needs to be removed during replication to allow proper action of PCNA at the replication forks. With regard to adjacent modification sites, a serine and a threonine, potential phosphorylation sites, are located next to lysine 56. Although H3S57 phosphorylation was reported to exist in mammals *in vivo* [46], no data on its appearance during cell cycle, on responsible enzymes and its function in mammals are available due to the lack of a specific antibody. One study,

applying yeast mutants proposes a potential functional interplay between H3K56 and S57 in replicative stress recovery and transcriptional elongation [46]. However, because H3S57ph has thus far not been identified in yeast *in vivo*, it is not possible to relate such observations to the mammalian system. Concerning putative H3K56me3-specific binding partners, we applied peptide pull-down experiments followed by MS identification of precipitated proteins (data not shown). Although we repeated such experiment many times, we were not able to consistently pull-down any candidates when compared to unmodified control peptide pull-downs. It is likely that H3K56me3 is not directly recognized by any “reader” protein but, instead functions indirectly by preventing acetylation of H3K56 and its associated signaling pathways. Alternatively, since H3K56me3 is localized in the α -N-helical region near the entry-exit sites of the DNA superhelix, it is possible that the correctly folded three-dimensional structure of this region (alone or in combination with DNA or other histones) is crucial for reader binding. Therefore, the use of peptides in such pull-down experiments will not suffice in reader binding. H3K56me3 histones or even nucleosomes containing this PTM will be needed for the identification of its potential reader(s) in the future.

Our finding that H3K56me3 constitutes another heterochromatin mark is in perfect agreement with previously published data, since H3K56 is monomethylated by G9A [12] that was initially described as a KMT responsible for H3K9me1 and H3K9me2 [47]. It is therefore plausible that H3K9me3-specific KMT(s) might also act on H3K56. We report here that the loss of Suv39h enzymes leads to diminished trimethylation of both H3K56 as well as H3K9. Based on our experimental set-up using Suv39h double-null cells, it is at the moment not possible to exclude that loss of H3K56me3 stems from an indirect effect. The chance of H3K9me3 influencing trimethylation of H3K56 by an, as yet, unknown mechanism, is conceivable albeit unlikely. Several observations argue for a direct enzymatic action of Suv39h on H3K56; the presence of a “K/S/T” motif in both regions and the fact that G9a, another H3K9-specific KMT is the responsible enzyme for H3K56me1 [12]. Therefore, we propose that Suv39h enzymes directly trimethylate H3K56 leading to a pericentric heterochromatin localization.

Although like both H3K9me3 and H4K20me3, H3K56me3 also constitutes a mark found in DAPI-dense regions, these modifications are not identical in their localization when looked at in greater detail. H3K9me3 stains telomeric repeats [32] and our results indicate that the majority of H3K56me3 does not. In contrast to H4K20me3, we found H3K56me3 in distinct chromatin foci during all mitotic phases, indicating that this novel mark is found in much more distinct heterochromatic loci. We plan to investigate this finding in future studies.

Besides our discovery of a novel histone modification site, our study raises one important question for many researchers dealing with PTMs and their biological functions. The finding that some enzymes might have several targets is supported by another recent study showing that pericentric localization of H3K64me3, another H3 core modification, also depends on Suv39h activity [48]. Therefore, the observed severe knock-down [14] and over-expression [49] phenotypes that were previously assigned to the sole loss or gain of H3K9me3, respectively, have to be reevaluated, since Suv39h enzymes affect not only H3K9, but also H3K64 as well as H3K56 trimethylation. It is possible that the assigned role of H3K9me3 in protecting genome stability and heterochromatic gene silencing [50] is in part shared by H3K56me3.

In agreement with the finding that H3K9-specific KMTs act on H3K56, we demonstrated a strong correlation between both

residues as to their KDM-specificity. Our study expands the list of known histone target residues of enzymes belonging to the JMJD2 family of demethylases since we could show that they act not only on H3K9me3 and, in some cases, H3K36me3 [38], but also on H3K56me3. Of particular interest is our characterization of mJmjd2E, a predicted pseudogene and its human homolog hKDM4DL, which codes for a, so far, uncharacterized protein. Because of hKDM4DL's strongest expression in human testis, it will be of great interest to determine if and why removal of the trimethylation of H3K9 and H3K56 is important in this special tissue. Perhaps it is crucial during the process of histone-protamine exchange and/or relaxation of pericentric heterochromatin in humans; a statement that will be difficult to address since the mouse enzyme is predicted to be a pseudogene and not expressed. hKDM4DL might, therefore, constitute a human or primate-specific protein. If so, then functional studies on hKDM4DL in testis will be hard, if not impossible to perform.

Our study clearly puts forward H3K56me3 as a novel modification, but we were unable to address its functional relevance. Usually, knock-down of the enzymes targeting the respective modification provide insights into its biological role; but since H3K9 and H3K56 methylations are affected by the same enzymatic machinery in mammals, we do not have any technical tool at hand to pinpoint, *in vivo*, one particular phenotype to H3K56me3. However, identification of genes that affect the two modifications slightly differently in the *C. elegans* intestine opens up the possibility of future functional studies, at least in this particular organism.

Interestingly, we identified MET-1, a H3K36 KMT homolog, as needed for wild type levels of both H3K9me3 and H3K56me3. It was previously suggested that H3K36 methylation might be a prerequisite for H3K9me3 in worms [41], and perhaps it is similarly required for H3K56me3 as well. Previous studies reported that H3K9me3 in the germline is independent of MET-2 [40], however H3K9me3 levels are significantly reduced in MET-2-depleted embryos [42]. These results indicate that different KMTs might be primarily used in different tissues. Consistent with this hypothesis, depletion of MET-2 and SET-25 significantly reduces H3K9me3 levels in embryos [42], and H3K56me3 levels in the intestine (this study), but their effect is less pronounced for H3K9me3 levels in the intestine. Future studies will be needed to reveal how the preference for different KMTs is regulated in different tissues.

We identified multiple KMTs required for normal levels of both H3K9me3 and H3K56me3. One possible explanation for the requirement of two or more methyltransferases is that one of these KMTs deposits mono- (and perhaps di-) methylation, while the second KMT deposits trimethylation, in a manner dependent on prior mono- or dimethylation. This model is similar to what was previously reported for MET-2 and SET-25 in embryos [42]. Alternative possibilities include indirect effects, perhaps involving non-histone targets for these proteins.

Early EM studies revealed that *C. elegans* embryos lack electron-dense material, classically associated with heterochromatin [51]. In addition, while in mammalian cells H3K9me3 co-localizes with DAPI-bright regions of pericentric heterochromatin, in *C. elegans*, H3K9me3 localizes to DAPI-faint regions [40], leading to the suggestion that *C. elegans* lacks heterochromatin or that heterochromatin is different in this species [40]. *C. elegans* chromosomes are holocentric, and in the absence of a localized centromere, the

phrase “pericentric” does not apply. Instead, the brightest foci of H3K9me3 in *C. elegans* nuclei associate with the nuclear lamina [42]. H3K9me3 is coincident with H3K27me3 and nuclear lamina protein LEM-1, all of which are enriched along chromosome arms [42,52]. Therefore, these regions most likely are similar to mammalian heterochromatin near the nuclear periphery, or lamin associated domains, LADs [53]. Our results show that H3K56me3 colocalizes with H3K9me3 in worms, suggesting that H3K56me3 likely marks these lamin associated domains.

In agreement with a specialized role of H3K56me3 in testis is the finding that sperm cells in *C. elegans* contain solely trimethylation of H3K56 but not of H3K9. It will be of interest to see if H3K56me3 has an evolutionary conserved role in germline development, although its functional implication might be different in different metazoans.

Supporting Information

Figure S1 Immunoblot peptide competition experiments to determine specificity of α H3K9me3 antibodies used in this study. α H3K9me3 antibodies from (A) Active Motif or (B) the Jenuwein laboratory [16] were pre-incubated with 2 μ g/ml competitor peptides before addition to immunoblots containing recombinant H3 protein (R) or acid extracted HeLa Kyoto histones (H) (top). Ponceau staining (bottom) serves as loading control. (TIF)

Figure S2 Members of the JMJD2 family of demethylases affect H3K56me3. (A) IF microscopy of HeLa Kyoto cells that were transfected with GFP-tagged mJmjd2a-d and human Jmjd2d homolog hKDM4 (green) and co-stained with α H3K56me3 antibody (red) and DAPI (DNA, blue). Arrows indicate transfected and GFP-positive cells. Scale bar = 10 μ m. See also Figure 4A for detailed PTM analysis of HeLa cells transfected with mJmjd2E-GFP. (B) IF microscopy of HeLa Kyoto cells that were transfected with mJmjd2E-GFP (green) and co-stained with various histone PTM-specific antibodies (red) and DAPI (DNA, blue). Arrows indicate transfected and GFP-positive cells. Scale bar = 10 μ m. See also Figure 4B that contains a listing of the results depicted here. (TIF)

Table S1 List of peptides used in peptide competition experiments. (DOCX)

Acknowledgments

We thank Christian Feller for providing histones from S2 cells. We are grateful to all members of the Hake laboratory, especially Raphael Kunisch, and we thank Lothar Schermelleh, Heinrich Leonhardt, and Boris Joffe for discussions and help.

Author Contributions

Conceived and designed the experiments: APMJ MW GC IS GS SBH. Performed the experiments: APMJ SB MH SP MW IS MS. Analyzed the data: APMJ SB MH SP MW GC IS GS SBH. Contributed reagents/materials/analysis tools: APMJ SB MH SP MW GC IS GS SBH. Wrote the paper: SBH.

References

- Kouzarides T (2007) Chromatin modifications and their function. *Cell* 128: 693–705.
- Campos EI, Reinberg D (2009) Histones: annotating chromatin. *Annu Rev Genet* 43: 559–599.

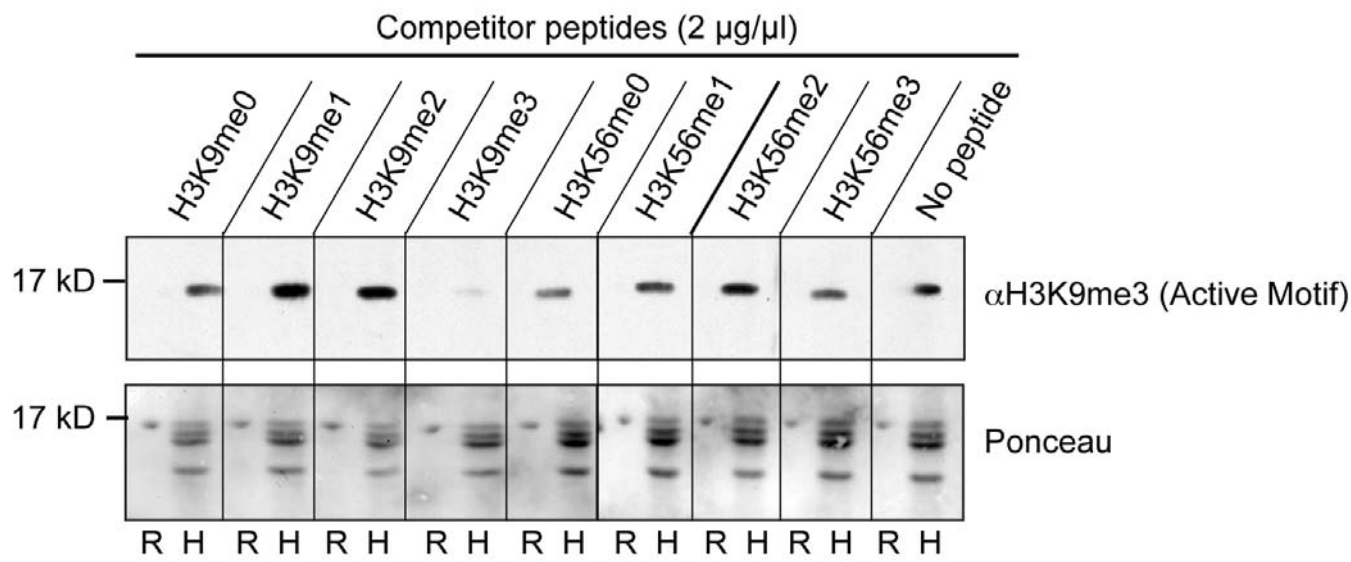
3. Strahl BD, Allis CD (2000) The language of covalent histone modifications. *Nature* 403: 41–45.
4. Xu F, Zhang K, Grunstein M (2005) Acetylation in histone H3 globular domain regulates gene expression in yeast. *Cell* 121: 375–385.
5. Xie W, Song C, Young NL, Sperling AS, Xu F, et al. (2009) Histone h3 lysine 56 acetylation is linked to the core transcriptional network in human embryonic stem cells. *Mol Cell* 33: 417–427.
6. Masumoto H, Hawke D, Kobayashi R, Verreault A (2005) A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response. *Nature* 436: 294–298.
7. Recht J, Tsubota T, Tanny JC, Diaz RL, Berger JM, et al. (2006) Histone chaperone Asf1 is required for histone H3 lysine 56 acetylation, a modification associated with S phase in mitosis and meiosis. *Proc Natl Acad Sci U S A* 103: 6988–6993.
8. Celic I, Masumoto H, Griffith WP, Meluh P, Cotter RJ, et al. (2006) The sirtuins hst3 and Hst4p preserve genome integrity by controlling histone h3 lysine 56 deacetylation. *Curr Biol* 16: 1280–1289.
9. Driscoll R, Hudson A, Jackson SP (2007) Yeast Rtt109 promotes genome stability by acetylating histone H3 on lysine 56. *Science* 315: 649–652.
10. Li Q, Zhou H, Wurtle H, Davies B, Horadzovsky B, et al. (2008) Acetylation of histone H3 lysine 56 regulates replication-coupled nucleosome assembly. *Cell* 134: 244–255.
11. Garcia BA, Hake SB, Diaz RL, Kauer M, Morris SA, et al. (2007) Organismal differences in post-translational modifications in histones H3 and H4. *J Biol Chem* 282: 7641–7655.
12. Yu Y, Song C, Zhang Q, Dimaggio PA, Garcia BA, et al. (2012) Histone H3 Lysine 56 Methylation Regulates DNA Replication through Its Interaction with PCNA. *Mol Cell*.
13. Neumann B, Held M, Liebel U, Erfle H, Rogers P, et al. (2006) High-throughput RNAi screening by time-lapse imaging of live human cells. *Nat Methods* 3: 385–390.
14. Peters AH, O'Carroll D, Scherthan H, Mechtler K, Sauer S, et al. (2001) Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* 107: 323–337.
15. Schotta G, Sengupta R, Kubicek S, Malin S, Kauer M, et al. (2008) A chromatin-wide transition to H4K20 monomethylation impairs genome integrity and programmed DNA rearrangements in the mouse. *Genes Dev* 22: 2048–2061.
16. Peters AH, Kubicek S, Mechtler K, O'Sullivan RJ, Derijck AA, et al. (2003) Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. *Mol Cell* 12: 1577–1589.
17. Shechter D, Dormann HL, Allis CD, Hake SB (2007) Extraction, purification and analysis of histones. *Nat Protoc* 2: 1445–1457.
18. Bonisch C, Schneider K, Punzeler S, Wiedemann SM, Bielmeier C, et al. (2012) H2A.Z.2.2 is an alternatively spliced histone H2A.Z variant that causes severe nucleosome destabilization. *Nucleic Acids Res*.
19. Hake SB, Garcia BA, Kauer M, Baker SP, Shabanowitz J, et al. (2005) Serine 31 phosphorylation of histone variant H3.3 is specific to regions bordering centromeres in metaphase chromosomes. *Proc Natl Acad Sci U S A* 102: 6344–6349.
20. Csankovszki G, McDonel P, Meyer BJ (2004) Recruitment and spreading of the C. elegans dosage compensation complex along X chromosomes. *Science* 303: 1182–1185.
21. Wells MB, Snyder MJ, Custer LM, Csankovszki G (2012) Caenorhabditis elegans dosage compensation regulates histone H4 chromatin state on X chromosomes. *Mol Cell Biol* 32: 1710–1719.
22. Wiedemann SM, Mildner SN, Bonisch C, Israel L, Maier A, et al. (2010) Identification and characterization of two novel primate-specific histone H3 variants, H3.X and H3.Y. *J Cell Biol* 190: 777–791.
23. Fodor BD, Kubicek S, Yonezawa M, O'Sullivan RJ, Sengupta R, et al. (2006) Jmjd2b antagonizes H3K9 trimethylation at pericentric heterochromatin in mammalian cells. *Genes Dev* 20: 1557–1562.
24. Kamath RS, Ahringer J (2003) Genome-wide RNAi screening in Caenorhabditis elegans. *Methods* 30: 313–321.
25. Ozdemir A, Masumoto H, Fitzjohn P, Verreault A, Logie C (2006) Histone H3 lysine 56 acetylation: a new twist in the chromosome cycle. *Cell Cycle* 5: 2602–2608.
26. Das C, Lucia MS, Hansen KC, Tyler JK (2009) CBP/p300-mediated acetylation of histone H3 on lysine 56. *Nature* 459: 113–117.
27. Vempati RK, Jayani RS, Notani D, Sengupta A, Galande S, et al. (2010) p300-mediated acetylation of histone H3 lysine 56 functions in DNA damage response in mammals. *J Biol Chem* 285: 28553–28564.
28. Yuan J, Pu M, Zhang Z, Lou Z (2009) Histone H3-K56 acetylation is important for genomic stability in mammals. *Cell Cycle* 8: 1747–1753.
29. Gu B, Watanabe K, Dai X (2012) Pygo2 regulates histone gene expression and H3 K56 acetylation in human mammary epithelial cells. *Cell Cycle* 11: 79–87.
30. Salic A, Mitchison TJ (2008) A chemical method for fast and sensitive detection of DNA synthesis in vivo. *Proc Natl Acad Sci U S A* 105: 2415–2420.
31. Schotta G, Lachner M, Sarma K, Ebert A, Sengupta R, et al. (2004) A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev* 18: 1251–1262.
32. Chadwick BP (2007) Variation in Xi chromatin organization and correlation of the H3K27me3 chromatin territories to transcribed sequences by microarray analysis. *Chromosoma* 116: 147–157.
33. Trievel RC (2004) Structure and function of histone methyltransferases. *Crit Rev Eukaryot Gene Expr* 14: 147–169.
34. Dambacher S, Hahn M, Schotta G (2010) Epigenetic regulation of development by histone lysine methylation. *Heredity (Edinb)* 105: 24–37.
35. Rea S, Eisenhaber F, O'Carroll D, Strahl BD, Sun ZW, et al. (2000) Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* 406: 593–599.
36. Tian X, Fang J (2007) Current perspectives on histone demethylases. *Acta Biochim Biophys Sin (Shanghai)* 39: 81–88.
37. Hou H, Yu H (2010) Structural insights into histone lysine demethylation. *Curr Opin Struct Biol* 20: 739–748.
38. Shin S, Janknecht R (2007) Diversity within the JMJD2 histone demethylase family. *Biochem Biophys Res Commun* 353: 973–977.
39. Whetstone JR, Nottke A, Lan F, Huarte M, Smolnikov S, et al. (2006) Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. *Cell* 125: 467–481.
40. Bessler JB, Andersen EC, Villeneuve AM (2010) Differential localization and independent acquisition of the H3K9me2 and H3K9me3 chromatin modifications in the Caenorhabditis elegans adult germ line. *PLoS Genet* 6: e1000830.
41. Andersen EC, Horvitz HR (2007) Two C. elegans histone methyltransferases repress lin-3 EGF transcription to inhibit vulval development. *Development* 134: 2991–2999.
42. Towbin BD, Gonzalez-Aguilera C, Sack R, Gaidatzis D, Kalck V, et al. (2012) Step-Wise Methylation of Histone H3K9 Positions Heterochromatin at the Nuclear Periphery. *Cell* 150: 934–947.
43. Tjeertes JV, Miller KM, Jackson SP (2009) Screen for DNA-damage-responsive histone modifications identifies H3K9Ac and H3K56Ac in human cells. *EMBO J* 28: 1878–1889.
44. Rufiange A, Jacques PE, Bhat W, Robert F, Nourani A (2007) Genome-wide replication-independent histone H3 exchange occurs predominantly at promoters and implicates H3 K56 acetylation and Asf1. *Mol Cell* 27: 393–405.
45. Williams SK, Truong D, Tyler JK (2008) Acetylation in the globular core of histone H3 on lysine-56 promotes chromatin disassembly during transcriptional activation. *Proc Natl Acad Sci U S A* 105: 9000–9005.
46. Aslam A, Logie C (2010) Histone H3 serine 57 and lysine 56 interplay in transcription elongation and recovery from S-phase stress. *PLoS One* 5: e10851.
47. Tachibana M, Sugimoto K, Nozaki M, Ueda J, Ohta T, et al. (2002) G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes Dev* 16: 1779–1791.
48. Daujat S, Weiss T, Mohn F, Lange UC, Ziegler-Birling C, et al. (2009) H3K64 trimethylation marks heterochromatin and is dynamically remodeled during developmental reprogramming. *Nat Struct Mol Biol* 16: 777–781.
49. Czvitkovich S, Sauer S, Peters AH, Deiner E, Wolf A, et al. (2001) Overexpression of the SUV39H1 histone methyltransferase induces altered proliferation and differentiation in transgenic mice. *Mech Dev* 107: 141–153.
50. Schotta G, Ebert A, Reuter G (2003) SU(VAR)3–9 is a conserved key function in heterochromatic gene silencing. *Genetica* 117: 149–158.
51. Leung B, Hermann GJ, Priess JR (1999) Organogenesis of the Caenorhabditis elegans intestine. *Dev Biol* 216: 114–134.
52. Gu SG, Fire A (2010) Partitioning the C. elegans genome by nucleosome modification, occupancy, and positioning. *Chromosoma* 119: 73–87.
53. Guen L, Pagie L, Brasset E, Meuleman W, Faza MB, et al. (2008) Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* 453: 948–951.

Supplementary Table 1 (Jack et al.)

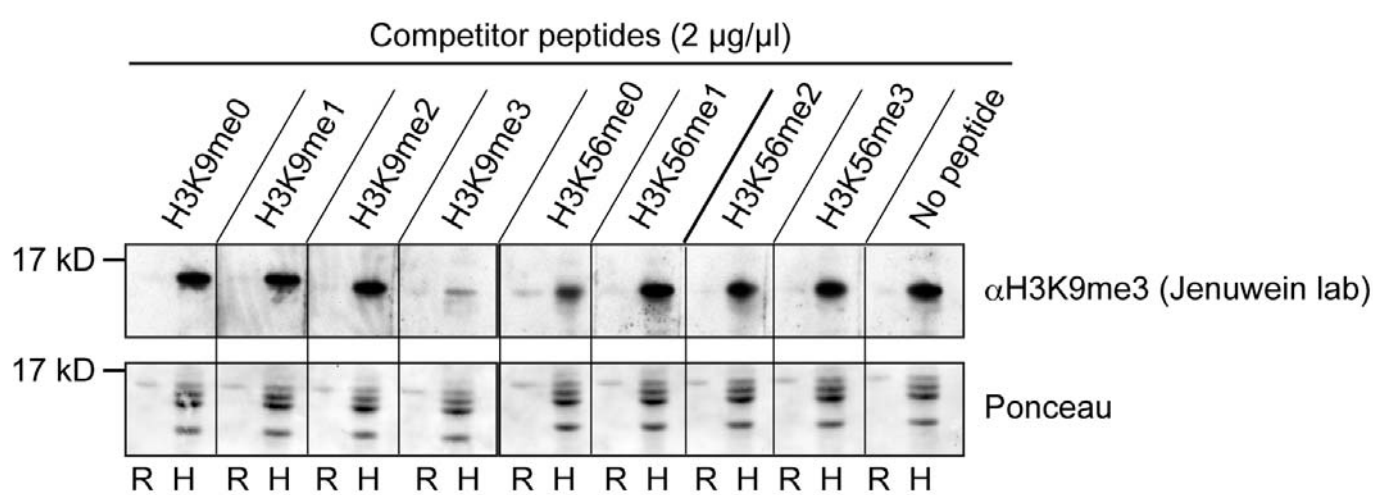
Peptide	Sequence
H3K56me0	VALREIRRYQKSTELLIRKL
H3K56me1	VALREIRRYQK(me1)STELLIRKL
H3K56me2	VALREIRRYQK(me2)STELLIRKL
H3K56me3	VALREIRRYQK(me3)STELLIRKL
H3K9me0	ARTKQTARKSTGGKAPRKQL
H3K9me1	ARTKQTARK(me1)STGGKAPRKQL
H3K9me2	ARTKQTARK(me2)STGGKAPRKQL
H3K9me3	ARTKQTARK(me3)STGGKAPRKQL
H3K4me3	ARTK(me3)QTARKSTGGKAPRKQL
H3K64me3	YQKSTELLIRK(me3)LPFQRLVRE
H3K79me3	LVREIAQDFK(me3)TDLRFQS

Supplementary Figure S1 (Jack et al.)

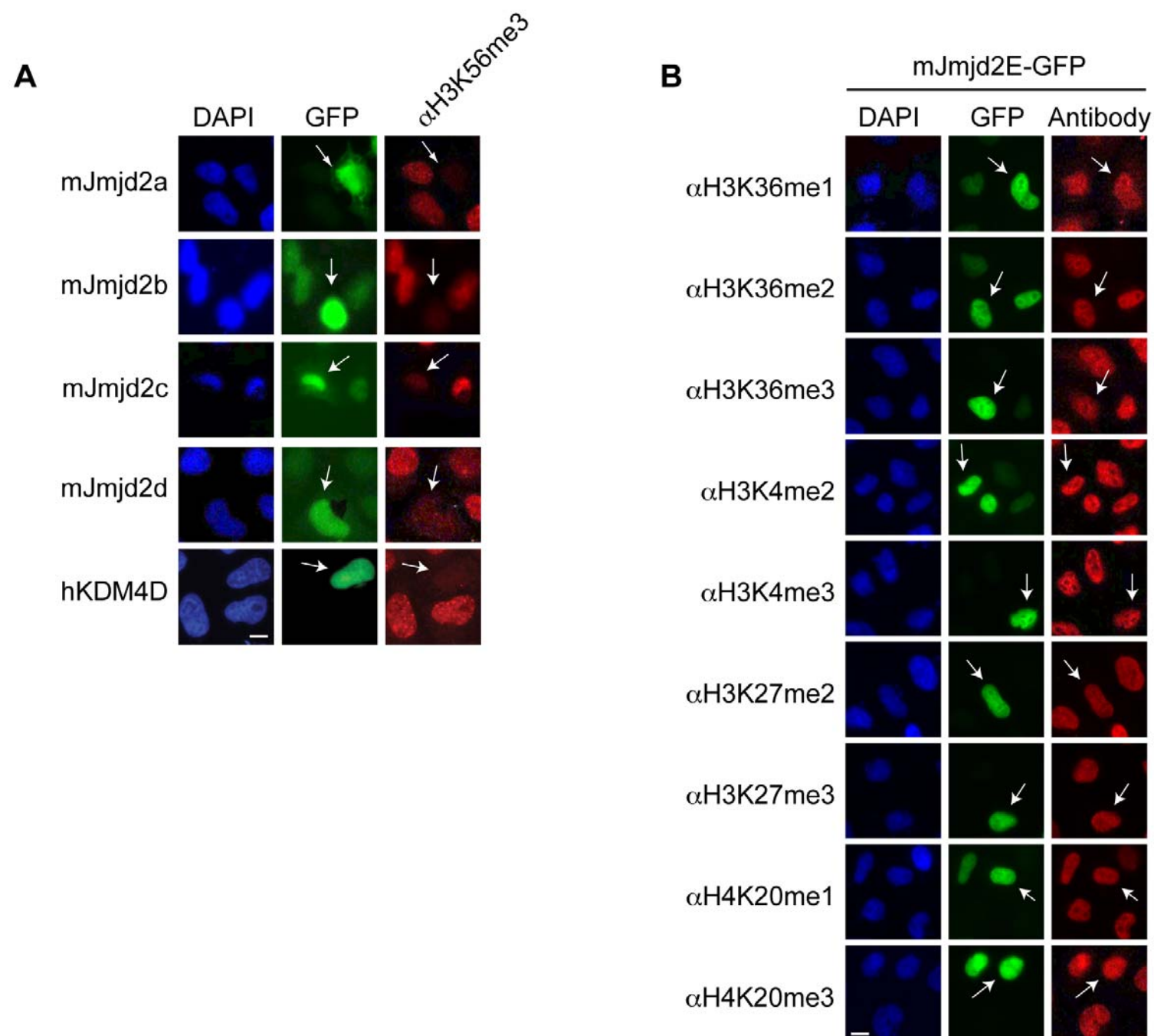
A



B



Supplementary Figure S2 (Jack et al.)



2.5 Suv4-20h2 mediates chromatin compaction and is important for cohesin recruitment to heterochromatin.

Article: Hahn et al. 2013, Genes & Development

Suv4-20h2 mediates chromatin compaction and is important for cohesin recruitment to heterochromatin

Matthias Hahn,^{1,2,9} Silvia Dambacher,^{1,2,9} Stanimir Dulev,^{1,2} Anastasia Yurievna Kuznetsova,³ Simon Eck,^{4,5} Stefan Wörz,^{4,5} Dennis Sadic,^{1,2} Maike Schulte,^{1,2} Jan-Philipp Mallm,^{5,6} Andreas Maiser,^{1,7} Pierre Debs,⁸ Harald von Melchner,⁸ Heinrich Leonhardt,^{1,7} Lothar Schermelleh,^{1,7,10} Karl Rohr,^{4,5} Karsten Rippe,^{5,6} Zuzana Storchova,³ and Gunnar Schotta^{1,2,11}

¹Munich Center for Integrated Protein Science (CiPS^M), Ludwig Maximilians University; ²Adolf-Butenandt-Institute, 80336 Munich, Germany; ³Max Planck Institute of Biochemistry, 82152 Martinsried, Germany; ⁴Department of Bioinformatics and Functional Genomics, Biomedical Computer Vision Group, Institut für Pharmazie und Molekulare Biotechnologie (IPMB), BioQuant, University of Heidelberg, 69120 Heidelberg, Germany; ⁵Deutsches Krebsforschungszentrum (DKFZ) Heidelberg, 69120 Heidelberg, Germany; ⁶Research Group Genome Organization and Function, BioQuant, 69120 Heidelberg, Germany; ⁷Department of Biology II, Biozentrum, 82152 Munich, Germany; ⁸Department of Molecular Hematology, University of Frankfurt Medical School, 60590 Frankfurt am Main, Germany

Cohesin plays an important role in chromatid cohesion and has additional functions in higher-order chromatin organization and in transcriptional regulation. The binding of cohesin to euchromatic regions is largely mediated by CTCF or the mediator complex. However, it is currently unknown how cohesin is recruited to pericentric heterochromatin in mammalian cells. Here we define the histone methyltransferase Suv4-20h2 as a major structural constituent of heterochromatin that mediates chromatin compaction and cohesin recruitment. Suv4-20h2 stably associates with pericentric heterochromatin through synergistic interactions with multiple heterochromatin protein 1 (HP1) molecules, resulting in compaction of heterochromatic regions. *Suv4-20h* mutant cells display an overall reduced chromatin compaction and an altered chromocenter organization in interphase referred to as “chromocenter scattering.” We found that *Suv4-20h*-deficient cells display chromosome segregation defects during mitosis that coincide with reduced sister chromatid cohesion. Notably, cohesin subunits interact with Suv4-20h2 both in vitro and in vivo. This interaction is necessary for cohesin binding to heterochromatin, as *Suv4-20h* mutant cells display substantially reduced cohesin levels at pericentric heterochromatin. This defect is most prominent in G0-phase cells, where cohesin is virtually lost from heterochromatin, suggesting that Suv4-20h2 is involved in the initial loading or maintenance of cohesin subunits. In summary, our data provide the first compelling evidence that Suv4-20h2 plays essential roles in regulating nuclear architecture and ensuring proper chromosome segregation.

[**Keywords:** Suv39h; Suv4-20h2; chromocenter clustering; chromosome segregation; cohesin; heterochromatin]

Supplemental material is available for this article.

Received November 15, 2012; revised version accepted March 26, 2013.

Pericentric heterochromatin in mammalian cells is formed from large arrays of noncoding satellite repeat sequences. Heterochromatic domains from different chromosomes can join into large clusters (so-called chromocenters) that can be visualized as DAPI-dense regions in interphase cells. Pericentric heterochromatin is largely transcriptionally inert

but serves important functions in ensuring genomic stability and accurate chromosome segregation (Peters et al. 2001; Ting et al. 2011; Zhu et al. 2011). Thus, dysregulation of heterochromatin organization leads to severe diseases and developmental defects (Hahn et al. 2010).

The major constitutive heterochromatin proteins that are thought to establish the proper chromatin structure at pericentric heterochromatin are heterochromatin protein 1 (HP1) isoforms and the histone methyltransferases Suv39h and Suv4-20h (Supplemental Fig. S1A). Suv39h enzymes induce histone H3 Lys 9 trimethylation (H3K9me3) at heterochromatic regions (Peters et al. 2001). This modification is recognized by HP1 molecules, which recruit

⁹These authors contributed equally to this work.

¹⁰Present address: Department of Biochemistry, University of Oxford, Oxford OX1 3QU, United Kingdom.

¹¹Corresponding author

E-mail gunnar.schotta@med.uni-muenchen.de

Article published online ahead of print. Article and publication date are online at <http://www.genesdev.org/cgi/doi/10.1101/gad.210377.112>.

Suv4-20h histone methyltransferases that subsequently establish H4K20me3 (Schotta et al. 2004, 2008). It is assumed that HP1 is a major regulator of heterochromatin organization and compaction. However, this view is challenged by the fact that HP1 proteins only transiently associate with heterochromatin (Cheutin et al. 2003). Therefore, other more constitutive components might contribute to regulating heterochromatin organization.

During mitosis, pericentric heterochromatin is important to facilitate sister chromatid cohesion. This is accomplished by cohesin complexes that connect sister chromatids at pericentric regions until anaphase onset (Salic et al. 2004; Tang et al. 2004). The cohesin ring consists of a Smc1–Smc3 dimer that is connected through the α -kleisin subunit Scc1/Rad21 (Nasmyth 2011). Cohesin is loaded onto chromatin in early G1 phase. Sites of cohesin loading can differ from regions with the highest cohesin enrichment, and therefore the cohesin rings are assumed to slide several kilobases from the loading site (Lengronne et al. 2004), where they might be fixed by interactions with other proteins, such as CTCF (Parelho et al. 2008; Rubio et al. 2008; Wendt et al. 2008). Due to the large size and compact structure of pericentric regions, there must be loading sites within heterochromatin, and it is therefore likely that core heterochromatin proteins assist in stable cohesin recruitment within these domains. In agreement with this hypothesis, cohesin is recruited to heterochromatic regions by Swi6/HP1 in fission yeast (Nonaka et al. 2002). However, HP1 proteins do not interact with cohesin in mammalian cells (Koch et al. 2008), and it is currently unknown how recruitment and maintenance of cohesin at pericentric heterochromatin is mediated.

In this study, we discovered an interaction between the histone methyltransferase Suv4-20h2 and cohesin. We found that Suv4-20h2 is required for cohesin recruitment to pericentric heterochromatin. *Suv4-20h*-deficient cells have strongly reduced cohesin levels at pericentric heterochromatin, resulting in chromosome segregation defects. Thus, our data demonstrate a novel role of the Suv39h–Suv4-20h pathway in cohesin recruitment to pericentric heterochromatin.

Results

Suv4-20h2 is a stable component of pericentric heterochromatin

We first set out to determine whether any of the heterochromatin core proteins could play structural roles in heterochromatin. We assessed their mobility in embryonic stem (ES) cells by fluorescence recovery after photobleaching (FRAP) analysis specifically at pericentric heterochromatin (Supplemental Fig. S1B). Consistent with previous reports (Cheutin et al. 2003), we found that HP1 α is a very mobile protein (Fig. 1A). This is a common feature of all HP1 variants, as the other isoforms—HP1 β and HP1 γ —also displayed the same fast recovery kinetics (K Rippe, pers. comm.). The histone methyltransferases Suv39h1 and Suv4-20h1 showed intermediate mobility.

Surprisingly, Suv39h2 and Suv4-20h2 stably associate with heterochromatin, as both proteins showed almost no recovery on the minute time scale (Fig. 1A). In addition, we determined the mobility parameters of the core heterochromatin proteins in another cell system: mouse embryonic fibroblast cells (MEFs). Interestingly, the recovery kinetics of all proteins were slower as compared with ES cells, which is in agreement with the hypothesis that heterochromatin is less plastic in differentiated cells (Meshorer et al. 2006). Similarly to ES cells, we found that Suv39h2 and Suv4-20h2 exhibit the strongest binding to heterochromatin (Fig. 1A). These experiments were complemented by fluorescence correlation spectroscopy (FCS) measurements of Suv4-20h2 at endogenous expression levels (Supplemental Fig. S2). For these assays, we generated a *Suv4-20h2*^{EGFP} knock-in ES cell line that expresses *Suv4-20h2* at nearly endogenous levels (Supplemental Fig. S2B) and displays clear enrichment of Suv4-20h2 and H4K20me3 at pericentric heterochromatin (Supplemental Fig. S2C). FCS mobility measurements confirm that at endogenous expression levels, the mobile Suv4-20h2 binds more tightly to chromatin than HP1, since its apparent diffusion coefficient, which includes the binding contribution, is reduced (Supplemental Fig. S2D). Furthermore, the immobile pool of Suv4-20h2 was ~ 10 times higher than that of HP1, as determined from continuous photobleaching experiments (Supplemental Fig. S2E). Taken together, our data suggest that the histone methyltransferases Suv39h2 and Suv4-20h2 could play structural roles in pericentric heterochromatin.

Stable binding of Suv4-20h2 is mediated through synergistic HP1 interactions

The stable binding of Suv4-20h2 to heterochromatin was also demonstrated in a previous study for human SUV4-20H2 (Souza et al. 2009), suggesting that this feature of Suv4-20h2 is evolutionarily conserved. However, as the bulk of HP1 only transiently associates with heterochromatin, it is surprising that Suv4-20h2 can stably bind heterochromatin. Therefore, we asked how this stable association of Suv4-20h2 is mediated. The C terminus of Suv4-20h2 is responsible for heterochromatin targeting (Schotta et al. 2004). In order to better define the targeting domain, we generated a panel of EGFP-tagged Suv4-20h2 truncation proteins as subfragments of the C terminus (Fig. 1B). Fragment M12, which comprises only 62 amino acids of the Suv4-20h2 C terminus, showed clear pericentric enrichment (Fig. 1B; Supplemental Fig. S2F). We further subdivided this fragment and found that two nonoverlapping truncations—M13 and M14—were both able to localize to pericentric heterochromatin as well (Fig. 1B; Supplemental Fig. S2F). These data indicate that at least two independent heterochromatin targeting modules exist in Suv4-20h2. We then asked whether the M12 region is really crucial for mediating pericentric recruitment of Suv4-20h2 and tested the localization of a Suv4-20h mutant protein lacking the M12 region. We did not detect enrichment of this mutant protein at pericentric

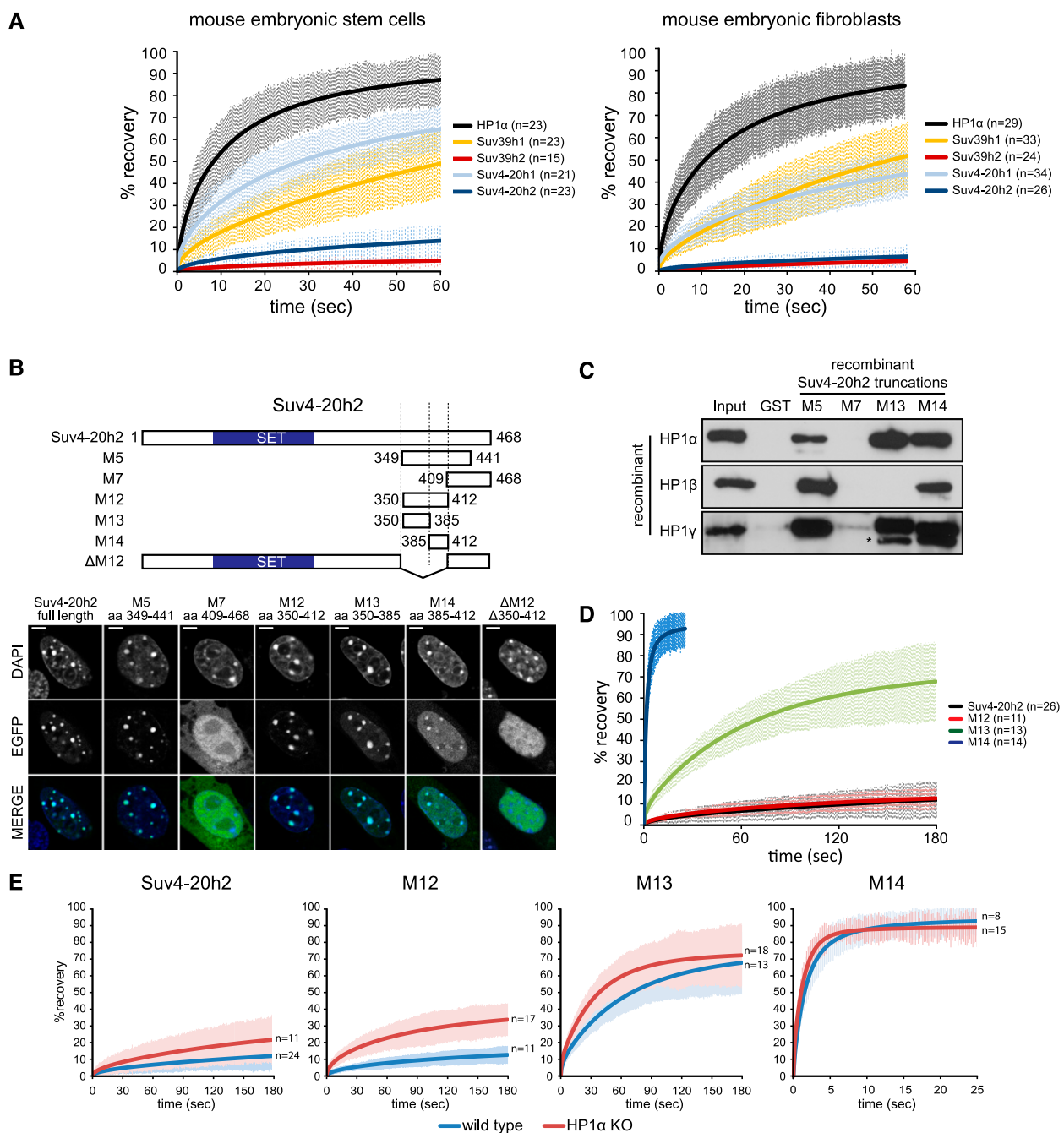


Figure 1. Suv4-20h2 is a stable component of pericentric heterochromatin. (A) FRAP analysis of core heterochromatin proteins. EGFP-tagged proteins were expressed in wild-type ES cells and MEFs. Laser bleaching and analysis of fluorescence recovery were carried out specifically at pericentric heterochromatin. FRAP recovery experiments for multiple cells (indicated in the legend) were averaged and fitted to a reaction diffusion model. Error bars represent the standard deviation of the FRAP measurements in each series. (B) Distinct regions in the C terminus mediate heterochromatin localization of Suv4-20h2. (Top panel) Schematic showing Suv4-20h2 truncation proteins that were expressed as EGFP fusion proteins in wild-type MEFs. (Bottom panel) Confocal sections of MEFs expressing EGFP-tagged Suv4-20h2 truncations. Bars, 5 μ m. (C) In vitro interaction test of Suv4-20h2 truncation proteins and HP1 isoforms. GST-tagged Suv4-20h2 truncation proteins were bound to GST beads and incubated with recombinant HP1 α , HP1 β , and HP1 γ , respectively. Bound proteins were separated on SDS page and probed with HP1 isoform-specific antibodies. The asterisk indicates cross-reacting bands with the HP1 γ antibody. (D) FRAP kinetics of Suv4-20h2 truncation proteins at pericentric heterochromatin. The C-terminal fragment M12 is as stable as the full-length Suv4-20h2 protein. Subfragments of M12 (M13 and M14) show a much faster recovery. FRAP recovery experiments for multiple cells (indicated in the legend) were averaged and fitted to a reaction diffusion model. Error bars represent the standard deviation of the FRAP measurements in each series. (E) Suv4-20h2 is more dynamic in HP1 α mutant cells. Suv4-20h2 full-length and truncation proteins were expressed in wild-type and HP1 α knockout cells. FRAP kinetics at heterochromatin of all Suv4-20h2 truncations are faster in HP1 α mutant cells. FRAP recovery experiments for multiple cells (indicated in the legend) were averaged and fitted to a reaction diffusion model. Error bars represent the standard deviation of the FRAP measurements in each series.

heterochromatin (Fig. 1B). Thus, we conclude that the region comprising amino acids 350–412 represents the heterochromatin targeting domain in Suv4-20h2.

Heterochromatin targeting of Suv4-20h2 depends on HP1 proteins (Schotta et al. 2004). We found that Suv4-20h2 can interact with all three mammalian HP1 isoforms in living cells (Supplemental Fig. S3). In vitro interaction assays revealed that all Suv4-20h2 truncations that localize to heterochromatin can directly interact with HP1 isoforms (Fig. 1C). Suv4-20h2 fragment M12 could not be generated as a soluble recombinant protein. However, as fragment M5 interacts with HP1 and fragment M7 does not show interaction with HP1, we conclude that the HP1 interaction domain resides within fragment M12. This is further supported by our finding that fragments M13 and M14, which are subregions of fragment M12, strongly interact with HP1 (Fig. 1C). Interestingly, HP1 α and HP1 γ appear as the major interaction partners with both M13 and M14, whereas HP1 β only interacts with fragment M14 (Fig. 1C). These data are in line with recent reports on HP1 interaction networks that demonstrate that HP1 α and HP1 γ share similar interaction partners that do not necessarily overlap with HP1 β interactors (Vermeulen et al. 2010). In summary, these data demonstrate that Suv4-20h2 has multiple independent HP1 interaction sites within its C terminus.

We further characterized the Suv4-20h2 truncations by measuring their mobility in heterochromatin using FRAP analysis. Full-length Suv4-20h2 stably binds to heterochromatin and showed a recovery of only ~10% at 3 min post-bleaching (Fig. 1D). The Suv4-20h2-M12 truncation behaved virtually identical to the full-length protein, indicating that this fragment comprises the essential domain that “clamps” Suv4-20h2 onto heterochromatin (Fig. 1D). We therefore refer to this region as the Suv4-20h2 “clamp domain.” Notably, the mobility of the M12 subfragments M13 and M14 was much higher. While the mobility of fragment M14 was in the range of HP1, M13 was more stably bound than HP1 (Fig. 1D), suggesting that distinct regions in Suv4-20h2 mediate the stable association with heterochromatin through interactions with HP1 proteins.

To test this hypothesis, we measured the mobility of the individual Suv4-20h truncations in fibroblasts with reduced HP1 levels. HP1 α knockout cells retained both H3K9me3 and H4K20me3 at pericentric heterochromatin, indicating that the recruitment of Suv39h and Suv4-20h enzymes to heterochromatin is not generally impaired (Supplemental Fig. S4A). Indeed, the localization of the Suv4-20h2 truncations was not altered in the absence of HP1 α (Supplemental Fig. S4B). However, the FRAP recovery kinetics of full-length Suv4-20h2 and truncations M12 and M13 were clearly enhanced (Fig. 1E). We did not observe faster recovery of fragment M14, as the FRAP kinetics of this fragment were already in the range of HP1 in wild-type cells. In summary, our data suggest that Suv4-20h2 is bound to heterochromatin through at least two independent interaction sites by HP1 proteins. As the individual interactions can only generate intermediate stability (M13 and M14), we postulate that

synergistic interaction with multiple HP1 proteins in the Suv4-20h2 clamp domain ensures its stable association with heterochromatin.

Suv4-20h2 regulates chromatin compaction

The tight interaction of Suv4-20h2 with heterochromatin suggests a direct involvement in regulating the structure of pericentric heterochromatin. To verify this hypothesis, we performed micrococcal nuclease (MNase) chromatin accessibility assays in wild-type, *Suv4-20h2* knockout, and *Suv4-20h* double-knockout ES cells (Schotta et al. 2008). Notably, we detected higher chromatin accessibility in both *Suv4-20h2* knockout and *Suv4-20h* double-knockout cells (Fig. 2A,B). The increased accessibility clearly involves heterochromatic regions, as demonstrated by Southern blotting of the digested DNA with major satellite probes (Fig. 2A,B). A similar increase in chromatin accessibility was also observed in *Suv4-20h*-deficient MEF cells, suggesting that Suv4-20h2 plays general roles in regulating chromatin structure (Supplemental Fig. S5). Importantly, re-expression of *Suv4-20h2*, which restores H4K20me3 (Nicetto et al. 2013), and even of the Suv4-20h2 clamp domain (M12) rescues the accessibility phenotype of *Suv4-20h* double-knockout cells (Supplemental Fig. S5), indicating that Suv4-20h2 can induce chromatin compaction through its clamp domain.

Superresolution three-dimensional (3D) structured illumination microscopy (3D-SIM) is a recently developed technique that allows imaging of subcellular structures below the optical diffraction limit (Schermerle et al. 2008). Using this technique, we detected a slightly reduced chromatin density in *Suv4-20h* double-knockout cells around the nuclear envelope (Fig. 2C; Supplemental Fig. S6). We quantified these data by measuring the DAPI intensity of the nuclear periphery compared with the overall DAPI intensity of the nuclei (Supplemental Fig. S6B). These data indicate that wild-type nuclei have a generally higher chromatin density in the nuclear periphery as compared with *Suv4-20h* double-knockout cells (Supplemental Fig. S6C). Interestingly, this reduction in peripheral heterochromatin correlates with changes in the distribution of nuclear pores, which are more stochastically arranged in *Suv4-20h* double-knockout nuclei (Fig. 2C). Strong overexpression of *Suv4-20h2* results in dramatic changes of the nuclear structure, most notably in increased chromatin compaction around chromocenters, nucleoli, and the nuclear envelope (Fig. 2C; Supplemental Fig. S6A). The increased density of the peripheral heterochromatin upon overexpression of *Suv4-20h2* is again reflected in the altered distribution of nuclear pores, which tend to be excluded from the highly compacted regions. In order to quantify the effect of Suv4-20h2 on nuclear pore organization, we developed an automatic 3D image analysis approach to measure nuclear pore parameters in our 3D-SIM images (Supplemental Fig. S6D). In agreement with the qualitative analysis of the nuclear pore staining, we found that the percentage of regions with low nuclear pore density (sparse nuclear pore regions) was

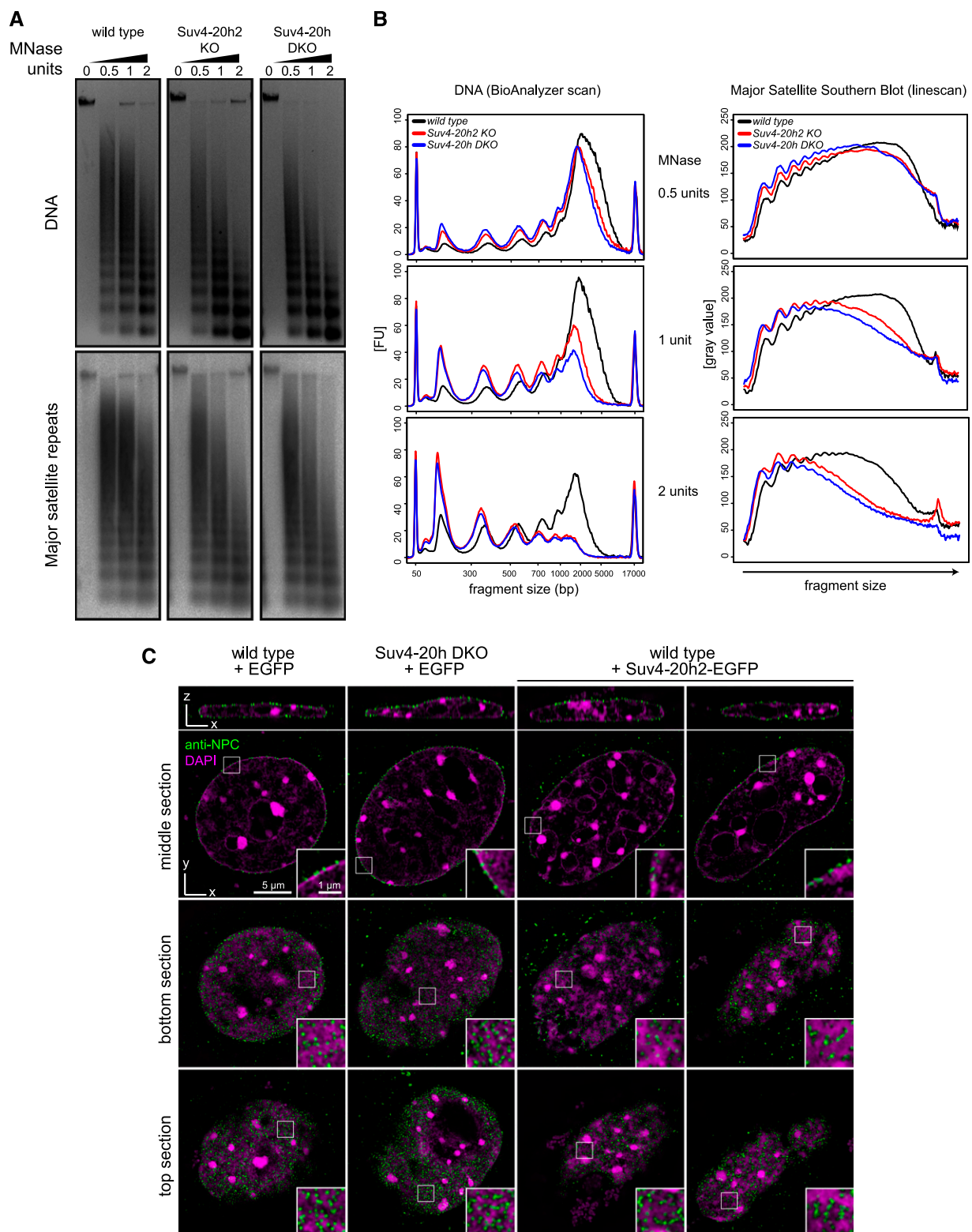


Figure 2. Suv4-20h2 mediates chromatin compaction. (A) MNase accessibility assay. Nuclei from wild-type, *Suv4-20h2* knockout, and *Suv4-20h* double-knockout ES cells were isolated and incubated with increasing amounts of MNase. (Top panel) The digested DNA was purified, separated on agarose gels, and stained with ethidium bromide. (Bottom panel) Heterochromatic regions were visualized by Southern blotting using a major satellite-specific probe. (B, left panel) Overlay of DNA electropherograms obtained from Bioanalyzer runs of the MNase-digested DNA from wild-type, *Suv4-20h2* knockout, and *Suv4-20h* double-knockout ES cells. The first and the last sharp peaks represent the markers of the Agilent DNA 12000 kit. (FU) Fluorescence units. (Right panel) Intensity line scan of major satellite Southern blots. Gray values correspond to signal intensities of the Southern blot. (C) 3D-SIM sections of wild-type and *Suv4-20h* double-knockout fibroblasts as well as wild-type cells overexpressing Suv4-20h2-EGFP. Immunofluorescence staining with an anti-NPC antibody marks nuclear pores. Insets show enlargements of the nuclear envelope. Quantifications of peripheral heterochromatin density and nuclear pore distribution are shown in Supplemental Figure S6.

reduced in *Suv4-20h* double-knockout cells and increased in cell overexpressing *Suv4-20h2* (Supplemental Fig. S6E).

We further asked which part of Suv4-20h2 would be responsible for the compaction phenotype. Overexpression of the N terminus containing the SET domain does not lead to obvious alterations of the nuclear structure (Supplemental Fig. S7). In contrast, the C terminus of Suv4-20h2 is very potent in inducing chromatin compaction (Supplemental Fig. S7). The Suv4-20h2 clamp domain does not show a compaction phenotype. Interestingly, these data show that Suv4-20h2 has important roles in regulating nuclear architecture, which is independent of its function as a histone methyltransferase.

Chromocenter scattering in *Suv4-20h*-deficient cells

Our data so far demonstrate that Suv4-20h2 is a major regulator at different levels of chromatin organization in vivo. On the one hand, the observed alterations in chromatin accessibility suggest that Suv4-20h2 regulates compaction of chromatin fibers. On the other hand, the structural changes in *Suv4-20h*-deficient and overexpressing cells indicate that Suv4-20h2 might even mediate long-range interactions between chromatin domains. We therefore investigated whether organization of pericentric heterochromatin is altered in *Suv4-20h* double-knockout cells. In mouse cells, heterochromatic regions of different chromosomes interact and form clusters, which are well discernible by DAPI staining. How this clustering is mediated is unknown. In order to visualize

pericentric heterochromatin very precisely, we performed fluorescence in situ hybridization (FISH) detection of major satellite repeats in G0-arrested cells to investigate the steady-state situation of heterochromatin in the absence of cell cycle perturbations (Fig. 3A). Stacks of confocal images were evaluated using an image segmentation procedure to detect and measure individual FISH foci (Supplemental Fig. S8). In wild-type cells, we detected ~25–30 foci (Fig. 3A,B), which is consistent with previous data on chromocenter clustering in mouse fibroblasts (Guenatri et al. 2004). *Suv4-20h* double-knockout cells, in contrast, showed chromocenter scattering, as we observed significantly more foci per nucleus (Fig. 3A,B; Supplemental Fig. S9). At the same time, the average focus volume was smaller in the mutant cells (Fig. 3B). Notably, re-expression of full-length *Suv4-20h2* in *Suv4-20h* double-knockout cells (DKO+FL) restores the altered nuclear organization of pericentric heterochromatin close to wild-type levels (Fig. 3B). Re-expression of the Suv4-20h2 clamp domain (M12) did not lead to a significant rescue (Fig. 3B; Supplemental Fig. S9), suggesting that although the clamp domain of Suv4-20h2 is necessary and sufficient to induce chromatin compaction as measured by MNase accessibility, additional features of the full-length protein are required to mediate long-range chromatin interactions like chromocenter clustering. Therefore, we asked whether chromocenter organization can be restored by a Suv4-20h2 mutant protein lacking methyltransferase activity (Nicetto et al. 2013). Re-expression of *Suv4-20h2*^{N182A,Y217A} (SET*) could not significantly reduce chromocenter

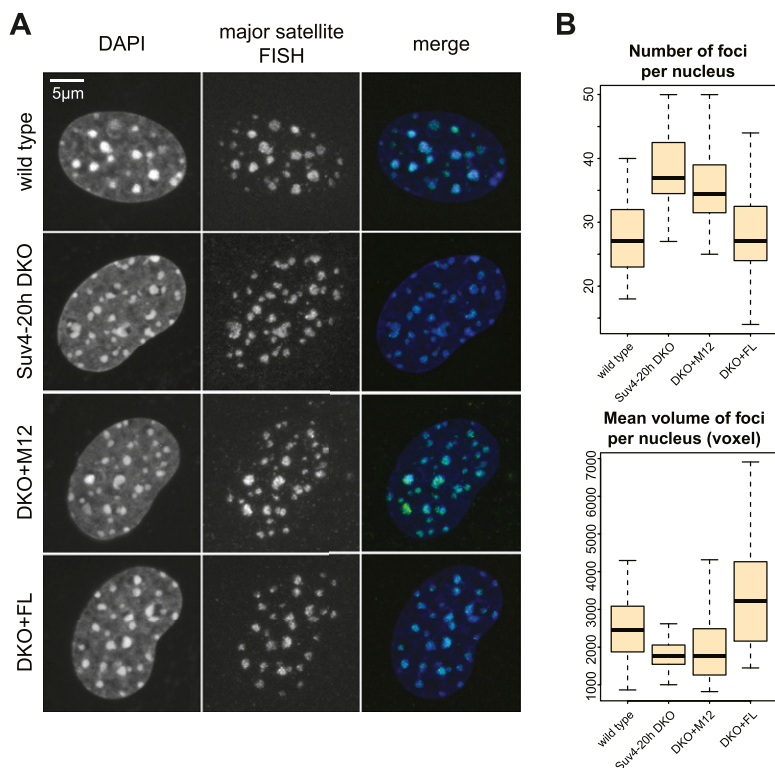


Figure 3. *Suv4-20h*-deficient cells display chromocenter scattering. (A) FISH analysis of major satellite repeats in G0-phase wild-type, *Suv4-20h* double-knockout, and *Suv4-20h2* rescue cells expressing Suv4-20h2 full-length protein (DKO+FL) or the Suv4-20h2 clamp domain (DKO+M12). (B) Quantification of major satellite FISH analyses. Chromocenter foci were counted and measured using an image segmentation analysis of confocal 3D stacks. Box plots showing the numbers of chromocenters and the chromocenter volume in wild-type ($n = 43$), *Suv4-20h* double-knockout ($n = 35$), DKO+FL ($n = 28$), and DKO+M12 ($n = 24$) nuclei. The mean number of foci per nucleus is significantly different between wild-type and *Suv4-20h* double-knockout as well as between *Suv4-20h* double-knockout and DKO+FL. Detailed statistical tests are provided in the Supplemental Material and in Supplemental Figure S9.

scattering (Supplemental Fig. S10), indicating that establishment of H4K20me3 is important for normal chromocenter organization.

Based on our data, we propose that Suv4-20h2 can mediate interactions between chromatin fibers by binding to H3K9me3-rich and HP1-rich domains. Loss of Suv4-20h enzymes would lead to a compromised heterochromatin organization, although H3K9me3 and HP1 are still present at heterochromatic regions.

Suv4-20h-deficient cells display chromosome segregation defects

Defects in heterochromatin organization often coincide with genomic instability. Therefore, we examined Suv4-20h mutant cells for chromosome segregation defects by analyzing mitotic figures. Significantly more mitotic abnormalities, such as anaphase bridges and lagging

chromosomes, could be observed in the *Suv4-20h* double-knockout cells (Fig. 4A). These defects can be rescued by expression of Suv4-20h2 full-length protein (DKO+FL) and even by expression of the Suv4-20h2 clamp domain (DKO+M12), demonstrating that Suv4-20h2 is important for proper chromosome segregation (Fig. 4A).

An analysis of different mitotic stages revealed that the segregation defects are not due to an altered progression through mitosis (Supplemental Fig. S11). Therefore, we tested whether key checkpoint proteins (AurB and Sgo1) or HP1 were altered in the absence of Suv4-20h enzymes. No obvious difference for either of these proteins was observed. However, we found that the centromere distance, as measured by the distance between adjacent CenpA foci, was wider in *Suv4-20h* double-knockout cells (Fig. 4B). Statistical evaluation confirmed an, on average, 100-nm increased centromere distance in mitotic spreads of *Suv4-20h* double-knockout cells (Fig. 4C).

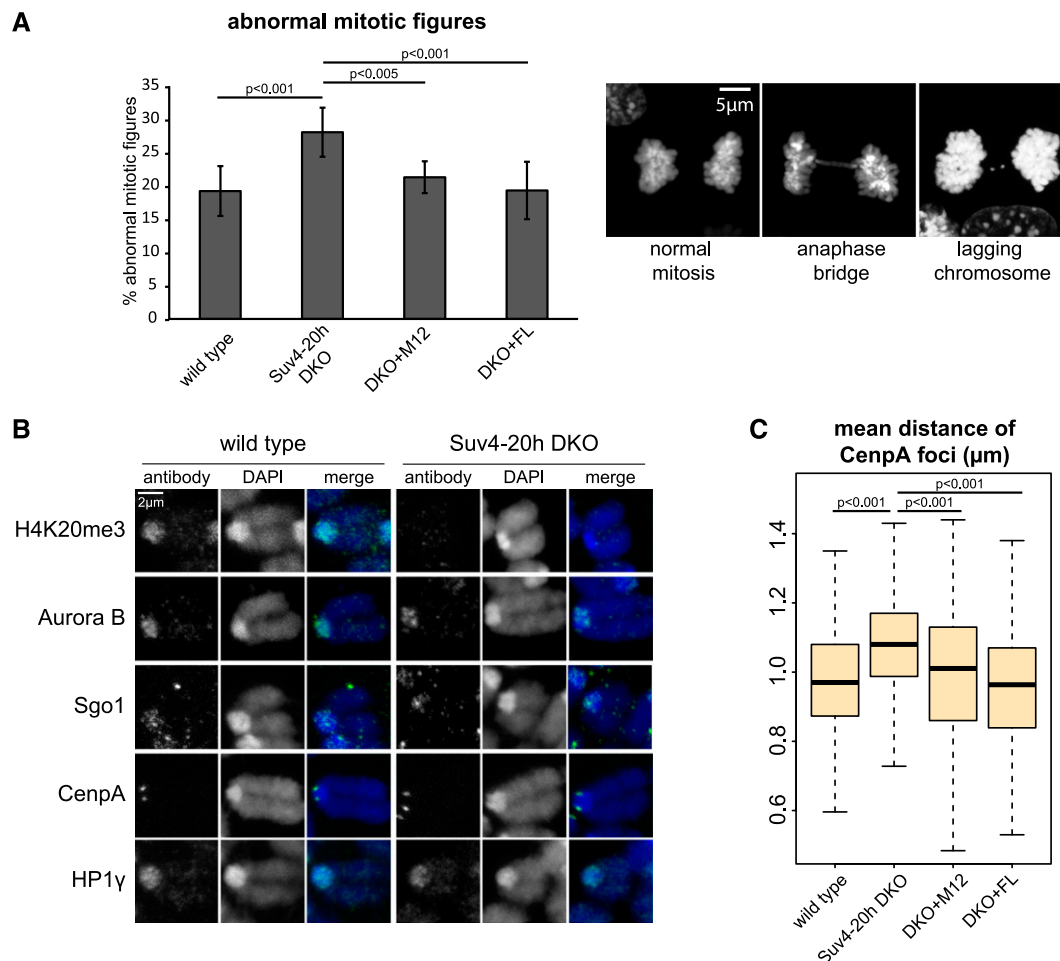


Figure 4. *Suv4-20h* mutants display chromosome segregation defects. (A) *Suv4-20h* mutant cells have mitotic chromosome segregation defects. Mitotic cells of wild-type ($n = 1497$), *Suv4-20h* double-knockout ($n = 1242$), and rescue MEFs (DKO+FL, $n = 768$; DKO+M12, $n = 687$) were assessed for anaphase bridges and lagging chromosomes (the right panel shows representative pictures of these defects). *P*-values of Student's *t*-test are indicated. Error bars correspond to standard deviation. (B) Mitotic marker analysis. Chromosome spreads of wild-type and *Suv4-20h* double-knockout cells were stained for specific markers that are known to be important for chromosome segregation. Representative pictures of mitotic chromosomes are shown. (C) Mean distance of CenpA foci in wild-type ($n = 458$), *Suv4-20h* double-knockout ($n = 348$), DKO+FL ($n = 242$), and DKO+M12 ($n = 318$) mitotic chromosomes. *P*-values of Student's *t*-test are indicated.

Re-expression of *Suv4-20h2* (DKO+FL) or of the Suv4-20h2 clamp domain (DKO+M12) lead to centromere distances comparable with wild-type cells (Fig. 4C). These data indicate that Suv4-20h2 is important for mediating normal sister chromatid cohesion. Reduced cohesion in *Suv4-20h* double-knockout cells in turn leads to mitotic abnormalities such as lagging chromosomes and anaphase bridges.

Suv4-20h2 interacts with cohesin

We thus hypothesized that in addition to having a structural role in heterochromatin, Suv4-20h2 may recruit additional proteins that facilitate sister chromatid cohesion. We performed GST pull-down assays to identify potential interaction partners of Suv4-20h2 using two different C-terminal truncation proteins (Supplemental Fig. S12A). The first region, M5, comprises the clamp domain, whereas the second fragment, M7, contains the very C terminus that does not localize to heterochromatin (see Fig. 1B). We could clearly precipitate all HP1 isoforms with fragment M5 (Fig. 5A), confirming the direct interaction of HP1 with the Suv4-20h2 clamp domain. Interestingly, additional proteins consistently found to coprecipitate with Suv4-20h2-M5 were subunits of the cohesin complex (Fig. 5A; Supplemental Table S1). The GST pull-down experiments were performed using sonicated nuclear extracts that contain genomic DNA and large chromatin fragments. In order to test whether cohesin indirectly associates with Suv4-20h2 through

DNA, we repeated the GST pull-down assays using extracts that were treated with benzonase to degrade all forms of RNA and DNA. Accordingly, extracts treated with benzonase were devoid of any contaminating DNA (Supplemental Fig. S12B). When using benzonase-treated extracts in GST pull-down assays, we clearly detected binding of both HP1 and cohesin subunits to Suv4-20h2-M5 (Fig. 5B). Similar results were obtained when extracts were treated with ethidium bromide, which disrupts DNA–protein interactions (Fig. 5B).

In order to verify that Suv4-20h2 interacts with cohesin subunits *in vivo*, we performed coimmunoprecipitation experiments using a *Suv4-20h2*^{HA-Flag} knock-in cell line. In this cell line, the endogenous *Suv4-20h2* locus is modified to produce a HA-3xFlag-tagged fusion protein, ensuring endogenous expression levels (Supplemental Fig. S12C). We developed a two-step extraction protocol to remove the bulk of cohesin through digestion of the DNA with benzonase (fraction I), followed by salt extraction of Suv4-20h2 and the remaining cohesin (fraction II). Flag affinity purification of Suv4-20h2 from fraction II extracts coprecipitated cohesin subunits (Fig. 5C). In the reverse experiment, cohesin subunit Smc1 could coprecipitate Suv4-20h2^{HA-Flag} (Fig. 5C). We did not detect interaction of unrelated proteins such as Suz12, confirming the specificity of the Suv4-20h2-cohesin interaction. In the pull-down experiments, we found that Suv4-20h2 fragment M5, containing the clamp domain, interacts with cohesin subunits. To test whether the clamp domain is important to mediate interaction with cohesin, we

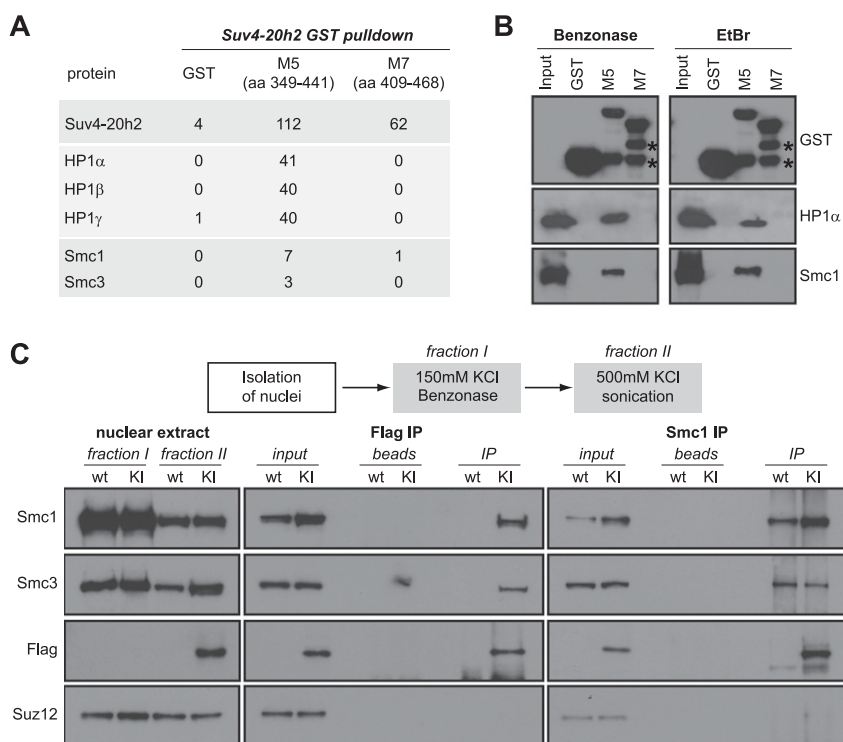


Figure 5. Suv4-20h2 interacts with cohesin subunits. (A) GST pull-down of Suv4-20h2 truncation proteins. Suv4-20h2 truncations—heterochromatin-associated (M5) and disperse nuclear (M7)—were expressed as recombinant GST-tagged proteins, incubated with nuclear extracts, and bound to affinity beads. Mass spectrometry identification of bound proteins revealed HP1 isoforms and two cohesin subunits to specifically interact with Suv4-20h2 fragment M5 (summary of triplicate experiments). The number of unique spectra that were identified for each protein in the mass spectrometry analysis is indicated (Scaffold analysis). The full list of identified proteins is shown in Supplemental Table S1. (B) GST pull-down experiments. Benzonase-treated nuclear extracts contain no detectable DNA contamination. Ethidium bromide treatment disrupts protein–DNA interactions. Western blots for GST pull-down experiments using recombinant GST, Suv4-20h2-M5, and Suv4-20h2-M7 were probed for GST, HP1 α , and Smc1. The Input lane contains the nuclear extract. Degradation products of the GST-tagged proteins are indicated by asterisks. (C) Suv4-20h2 and cohesin interact *in vivo*. Fractionated nuclear extracts were prepared from wild-type (wt) and *Suv4-20h2*^{HA-Flag} knock-in (KI) ES cells. Suv4-20h2 and Smc1 were precipitated from fraction II extracts using Flag and Smc1 antibodies, respectively. Bound proteins were visualized by Western blotting using Flag, Smc1, Smc3, and Suz12 antibodies.

and *Suv4-20h2*^{HA-Flag} knock-in (KI) ES cells. Suv4-20h2 and Smc1 were precipitated from fraction II extracts using Flag and Smc1 antibodies, respectively. Bound proteins were visualized by Western blotting using Flag, Smc1, Smc3, and Suz12 antibodies.

performed immunoprecipitation experiments in cells expressing full-length *Suv4-20h2* or *Suv4-20h2*- Δ M12, which lacks the clamp domain. We detected interaction of cohesin with full-length *Suv4-20h2* but not with *Suv4-20h2*- Δ M12 (Supplemental Fig. S12D), suggesting that the clamp domain is important to facilitate interaction with cohesin.

Reduced heterochromatin-associated cohesin in *Suv4-20h*-deficient cells

To test whether *Suv4-20h2* plays a role in recruiting cohesin to heterochromatin, we performed chromatin

immunoprecipitation (ChIP) experiments for cohesin subunits and tested their enrichment at major satellite repeats and control regions outside of pericentric heterochromatin. In wild-type cells, heterochromatin is enriched for repressive histone modifications (H3K9me3 and H4K20me3) and cohesin subunits (Fig. 6A; Supplemental Fig. S13). Strikingly, in *Suv4-20h* double-knockout cells, cohesin subunits were strongly reduced at major satellite repeats (Fig. 6A; Supplemental Fig. S13) but not at the control regions where cohesin is recruited through different mechanisms (Parelho et al. 2008; Rubio et al. 2008; Wendt et al. 2008). We observed reduced heterochromatin-associated cohesin in two independent *Suv4-20h2*

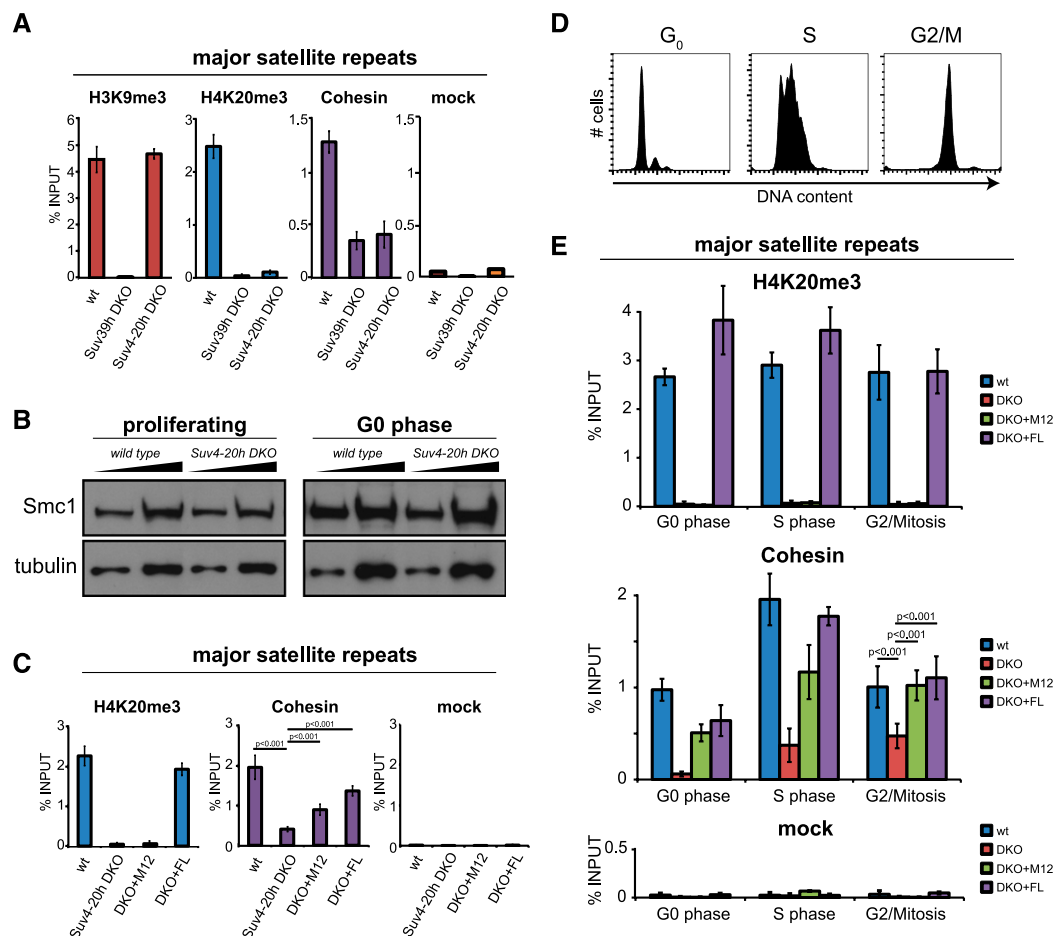


Figure 6. *Suv4-20h2* is important for cohesin recruitment to pericentric heterochromatin. (A) *Suv39h* double-knockout and *Suv4-20h* double-knockout cells display reduced cohesin levels at pericentric heterochromatin. ChIP analysis of histone modifications and the cohesin subunits Smc3 in wild-type, *Suv39h* double-knockout, and *Suv4-20h* double-knockout cells at major satellite repeats. Merge of two experiments in two independent cell lines with three technical replicates each. Error bars represent standard deviation over all experiments. (B) Cohesin levels are not generally impaired in *Suv4-20h* double-knockout cells. Protein extracts of proliferating (left panel) or G₀-arrested (right panel) wild-type and *Suv4-20h* double-knockout cells were tested for cohesin levels by Western blot for Smc1 (cohesin) and tubulin (loading control). Two different amounts of cell extracts were loaded. (C) Cohesin establishment can be mediated by the *Suv4-20h2* clamp domain. ChIP analysis of H4K20me3 and cohesin (Smc3) at major satellite repeats. Expression of full-length *Suv4-20h2* (DKO+FL) restores H4K20me3 and cohesin in *Suv4-20h* double-knockout cells. Cohesin is partially rescued by expression of the *Suv4-20h2* clamp domain (DKO+M12). Merge of two independent experiments with three technical replicates. Error bars represent standard deviation over all experiments. (D) FACS analysis of wild-type cells in G₀, S, and M phase. Cells were isolated as described and stained with propidium iodide prior to FACS analysis. (E) *Suv4-20h* enzymes are essential for cohesin recruitment to heterochromatin. Wild-type, *Suv4-20h* double-knockout, and rescue cell lines (DKO+M12 and DKO+FL) were harvested in G₀, S, and M phase. ChIP analysis for H3K9me3, H4K20me3, and cohesin (Smc3) at major satellite repeats was performed in two independent experiments with three technical replicates. Error bars represent standard deviation over all experiments.

knockout and two independent *Suv4-20h* double-knockout cell lines (Supplemental Fig. S13). Cohesin levels are not generally reduced in *Suv4-20h* double-knockout cells (Fig. 6B), suggesting specific defects in recruitment of cohesin to heterochromatin. Association of Suv4-20h2 with pericentric heterochromatin depends on Suv39h-mediated H3K9me3 and HP1 association. Therefore, we tested cohesin levels in *Suv39h* double-knockout cells that have lost both H3K9me3 and H4K20me3 at pericentric heterochromatin (Fig. 6A). Importantly, reduced cohesin levels could also be detected in two independent *Suv39h* double-knockout cell lines (Fig. 6A; Supplemental Fig. S14), demonstrating that the Suv39h–Suv4-20h pathway is important for cohesin recruitment to heterochromatin.

We then asked whether Suv4-20h2 is responsible for cohesin recruitment to heterochromatin by rescue experiments in *Suv4-20h* double-knockout cells. Expression of full-length *Suv4-20h2* (DKO+FL) could rescue both cohesin and H4K20me3 at major satellite repeats (Fig. 6C). Interestingly, the *Suv4-20h2* clamp domain (DKO+M12) could also increase cohesin levels at heterochromatin, although this truncation protein lacks the SET domain and thus was unable to restore H4K20me3 (Fig. 6C). Cohesin at control regions was not affected in the different rescue experiments (Supplemental Fig. S15). We conclude that Suv4-20h2 plays an important role in recruiting cohesin to heterochromatic regions through interactions with cohesin subunits.

Suv4-20h2 is important for the initial recruitment of cohesin to pericentric heterochromatin

Finally, we investigated whether Suv4-20h2 is involved in the initial recruitment/maintenance of cohesin at heterochromatin. Loading of cohesin onto chromatin occurs in early G1 phase (Watrin et al. 2006). In the subsequent cell cycle stages, cohesin is maintained on the chromosomes until mitosis. During metaphase, cohesin is removed from chromosome arms by proteolytic cleavage (Hauf et al. 2001). Only cohesin at pericentric heterochromatin is maintained until the onset of anaphase to ensure sister chromatid cohesion (Salic et al. 2004; Tang et al. 2004). In order to investigate whether Suv4-20h2 is implicated in cohesin loading or maintenance, we analyzed cohesin levels in different cell cycle stages. We synchronized cells in G0, S, and G2/M phase (Fig. 6D; Supplemental Fig. S16) and prepared chromatin for ChIP analysis. We found that cohesin levels varied only slightly over the cell cycle in wild-type cells (Fig. 6E). In contrast, in *Suv4-20h* double-knockout cells, cohesin was basically absent from heterochromatic regions in G0 phase, suggesting that Suv4-20h enzymes are crucial for loading and/or maintaining cohesin at heterochromatin. At later stages in the cell cycle (S and G2/M phase), cohesin levels increased (Fig. 6E), indicating that additional, Suv4-20h-independent recruitment mechanisms exist during S phase. The loss of heterochromatin-associated cohesin in *Suv4-20h* double-knockout cells could be rescued by expressing the clamp domain (M12) or full-length Suv4-20h2 (Fig. 6E),

whereas cohesin recruitment at unrelated control regions was not affected in the rescue cells (Supplemental Fig. S17). Thus, our data demonstrate that Suv4-20h2 is an important factor for cohesin loading onto heterochromatin.

Discussion

Heterochromatin is a very abundant chromatin state in mammalian cells, as >30% of the mammalian genome is composed of repetitive sequences that need to be silenced. Heterochromatin is therefore characterized by low transcriptional activity and features a less accessible chromatin structure. These special properties of heterochromatin are established in a step-wise manner. First, specific proteins need to recognize heterochromatic sequences (Bulut-Karslioglu et al. 2012). Second, heterochromatin becomes marked with distinct histone modifications that recruit additional binding proteins to ultimately mediate the special properties of this chromatin state. In vivo, heterochromatin is not a stiff entity, but rather a dynamic equilibrium of a protein interaction network. Thus, it can both reduce chromatin accessibility under steady-state conditions and open up chromatin structure in response to specific stimuli; for example, DNA damage (Goodarzi et al. 2008). Here we demonstrate that Suv4-20h2 is an important constituent of this interaction network for regulating chromatin accessibility and long-range chromatin interactions. In ES cells and fibroblasts, these functions do not seem to be shared with Suv4-20h1, which features a different heterochromatin targeting domain that is not homologous to the Suv4-20h2 clamp domain and does not bind so tightly to heterochromatin (Supplemental Fig. S18A).

How is Suv4-20h2 integrated into this network and how can it affect heterochromatin structure? Heterochromatin displays abundant H3K9me3, which provides a binding interface for HP1 proteins. Although HP1 shows only a weak affinity to H3K9me3 in vitro, the high density of H3K9me3 can lead to a high HP1 abundance in heterochromatin. We found that Suv4-20h2 features multiple HP1 interaction sites in its clamp domain, which can explain its stable association with regions that feature high HP1 concentrations. The interaction with multiple HP1 proteins raises an interesting hypothesis that Suv4-20h2 might bridge H3K9me3- and HP1-rich chromatin fibers to render chromatin less accessible. In agreement with this view, we found that chromatin in cells lacking *Suv4-20h2* is indeed more accessible to nucleases. Notably, the level of Suv4-20h2 can alter the balance within the heterochromatin interaction network and may therefore modulate the functions of its interacting proteins, like HP1. Another protein that may be affected by Suv4-20h2 is the linker histone H1. This is inferred by the observed alterations in the nucleosome repeat length in *Suv4-20h* double-knockout cells (Fig. 2B), which is a typical feature of cells with reduced levels of histone H1 (Fan et al. 2005). The current knowledge of the topology of the heterochromatin interaction network is insufficient to distinguish direct effects from indirect effects. However, our data demonstrate that Suv4-20h2

is a central node within this network, and therefore changes of Suv4-20h2 levels lead to drastic alterations in nuclear architecture.

Interestingly, the clamp domain of Suv4-20h2 is necessary and sufficient for chromatin compaction, as *Suv4-20h* double-knockout cells expressing this domain display normal chromatin accessibility. However, the clamp domain is not sufficient to mediate long-range chromatin interactions and condensation of large chromatin domains. Therefore, other domains in Suv4-20h2, such as an active SET domain, or recruitment of additional proteins are required for these functions. An intriguing possibility is that heterochromatin-associated cohesin, which is recruited by Suv4-20h2, is implicated in these long-range interactions (Hadjur et al. 2009; Kagey et al. 2010; Degner et al. 2011). Recent evidence suggests that cohesin may be involved in regulating the condensation of heterochromatic domains. Cohesin is present in a balance between a chromatin-associated form and a free form that is adjusted by a loading–releasing cycle (Kueng et al. 2006; Gause et al. 2010). Perturbation of this balance by knockdown of the cohesin-releasing factor WAPL leads to increased cohesin association with chromatin and a dramatic chromatin compaction phenotype (Seitan and Merkenschlager 2012). However, we could not detect significant perturbations in chromocenter organization in cells lacking the cohesin subunit Scc1/Rad21 (Supplemental Fig. S19). Therefore, we expect that other proteins that interact with Suv4-20h2 outside of the clamp domain are necessary to mediate chromocenter clustering.

Although cohesin appears to be dispensable for nuclear architecture in interphase cells, heterochromatin-associated cohesin is absolutely crucial for sister chromatid cohesion during mitosis. How cohesin is loaded onto chromatin and how the cohesin ring stabilizes interactions between chromatids are currently unclear (Nasmyth 2011). We found that the Suv39h–Suv4-20h pathway is important for the loading of cohesin to pericentric heterochromatin (Fig. 7). Our data are consistent with a study demonstrating that *Suv39h*-deficient cells display defects in sister chromatid cohesion and chromosome segregation (Koch et al. 2008). Although Koch et al. (2008) could not demonstrate that cohesin levels in *Suv39h* double-knockout cells were reduced, this may be due to their semiquantitative ChIP analysis, which is not suitable for highly repetitive sequences. We clearly detected reduced cohesin levels at major satellite repeats in different *Suv39h* double-knockout, *Suv4-20h* double-knockout, and *Suv4-20h2* knockout cell lines. In none of these cell lines, cohesin was completely lost from heterochromatin, suggesting additional loading pathways. Based on our data, we conclude that the primary function of Suv39h enzymes is to prepare pericentric heterochromatin for HP1 recruitment, which is a prerequisite for stable association of Suv4-20h2. Cohesion recruitment is then facilitated through interactions of cohesin subunits with Suv4-20h2 and possibly additional factors. Suv4-20h1 does not interact with cohesin and is therefore unlikely to contribute to cohesin recruitment (Supplemental Fig. S18C). Interestingly, cohesin recruitment can be mediated

at least partially through the Suv4-20h2 clamp domain. As this domain can interact with cohesin and also induce chromatin compaction, it is not possible to distinguish which of the two functions is more relevant for cohesin recruitment to heterochromatin.

The cell cycle analysis revealed that Suv4-20h2 is essential for cohesin loading in G0 phase. During later stages of the cell cycle, Suv4-20h2-independent pathways can recruit cohesin to heterochromatin. This would explain why mitosis is not completely blocked and why comparably mild mitotic phenotypes were detected in *Suv4-20h*-deficient cells. Importantly, our data are consistent with other reports that demonstrate that the fidelity of mitosis is already compromised when heterochromatin-associated cohesin levels are only reduced (Eckert et al. 2007). A low level of mitotic defects can therefore ensure survival of the cells but contributes to genomic instability.

Dysregulation of pericentric heterochromatin has been suggested to play important roles in cancer development and progression (Hahn et al. 2010; Zhu et al. 2011). Cancer cells are frequently characterized by genomic instability and cohesion defects (Thompson et al. 2010), but the mechanisms are still poorly understood. Previous analyses have shown that a reduced level of H4K20me3, which is an indirect measure for the presence of Suv4-20h enzymes, characterizes different human tumors (Fraga et al. 2005). In light of our data, we hypothesize that Suv4-20h2 might be an important cohesin recruitment factor in human cells. Its dysregulation could lead to reduced levels of heterochromatin-associated cohesin, which in turn contributes to the genomic instability that is characteristic of many human tumors and could explain the negative survival prognosis with tumors that have low H4K20me3 levels (Van Den Broeck et al. 2008; Schneider et al. 2011).

Materials and methods

FRAP

FRAP measurements were performed in MEF and ES cell lines after transient expression of EGFP-tagged proteins. Analysis of the recovery curves of the intensity integrated over the region of interest (ROI) was done using FREDIS software (Muller et al. 2009). For Figure 1, the averaged recovery curves were fitted to a diffusion model or a reaction diffusion model that incorporates both diffusion and binding processes.

HP1 interaction test

GST fusion proteins of Suv4-20h2, HP1 α , HP1 β , and HP1 γ were expressed in *Escherichia coli* and purified using glutathione-S-sepharose beads (Amersham Biosciences). The GST tag was subsequently removed using PreScission protease. Interaction tests were performed by incubating 5 μ g of GST-Suv4-20h2 fusion protein with 5 μ g of either HP1 α , HP1 β , or HP1 γ for 1 h at room temperature on a rotating wheel.

Chromatin accessibility assay

Nuclei of wild-type, *Suv4-20h2* knockout, and *Suv4-20h2* double-knockout ES cells and wild-type, *Suv4-20h* double-knockout,

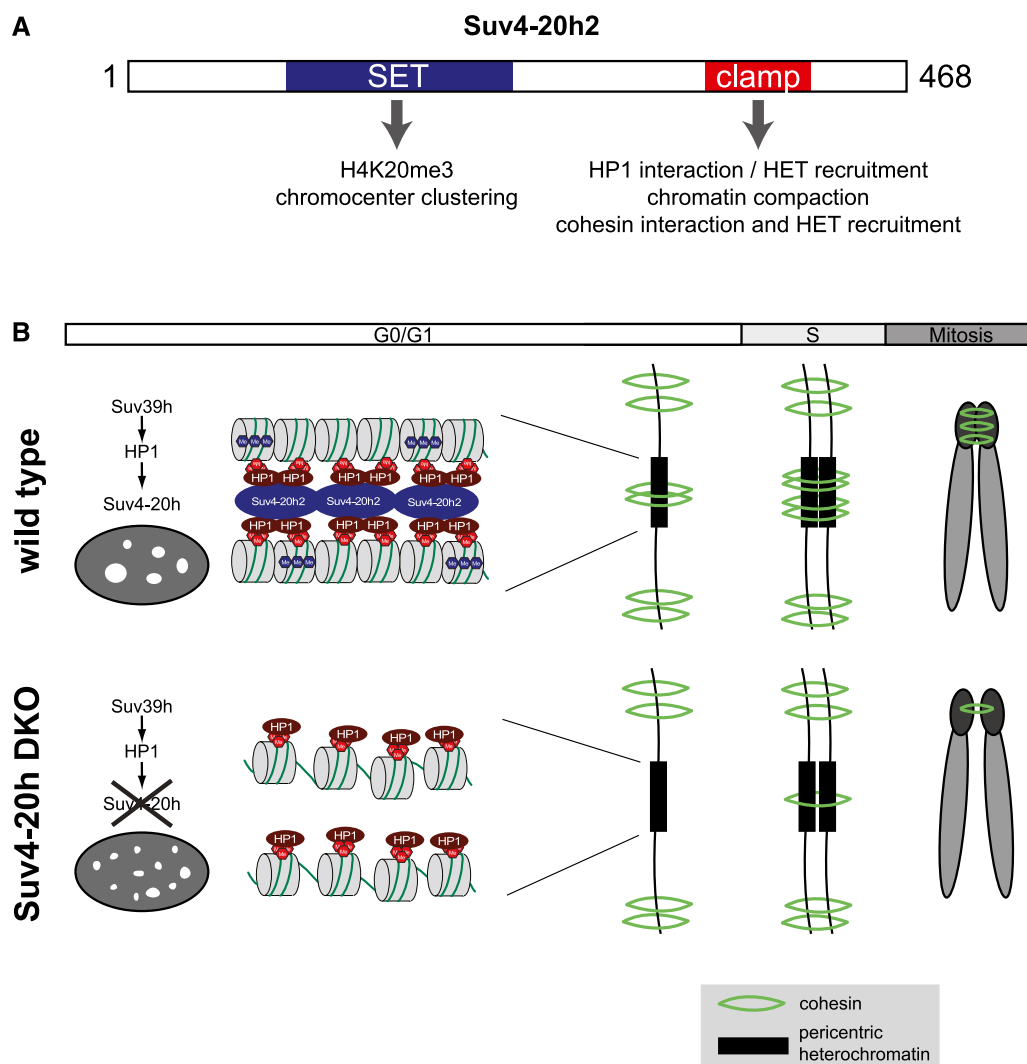


Figure 7. Model for Suv4-20h2-mediated chromatin compaction. (A) Functional domains in Suv4-20h2. The SET domain is the catalytic domain that induces H4K20me3. Establishing this modification is important to mediate chromocenter clustering. The clamp domain is essential to recruit Suv4-20h2 to pericentric heterochromatin. The clamp domain also mediates the Suv4-20h2-cohesin interaction and thus the recruitment of cohesin to heterochromatin. (B) Model showing the role of the Suv39h–HP1–Suv4-20h pathway for chromatin compaction and cohesin recruitment to pericentric heterochromatin. Based on our data, we propose that Suv4-20h2 can mediate interactions between chromatin fibers by binding to H3K9me3-rich and HP1-rich domains. The loss of Suv4-20h enzymes leads to a compromised heterochromatin organization (reduced chromatin compaction and chromocenter scattering) although H3K9me3 and HP1 are still present at heterochromatic regions. In G0/G1 cells, the interaction with Suv4-20h2 is important for cohesin recruitment to pericentric heterochromatin. Cohesin is not completely lost from mitotic chromosomes due to alternative, Suv4-20h-independent loading pathways during S/G2 phase.

DKO+FL, and DKO+M12 fibroblast cell lines were digested with different amounts of MNase (Sigma) according to Gilbert et al. (2007). Genomic DNA was purified and separated on a 1% agarose gel or loaded onto a DNA LabChip (Agilent Technologies).

3D-SIM

Superresolution 3D-SIM of immunofluorescently labeled and DAPI-stained nuclei was performed as previously described (Schermetleh et al. 2008) on a DeltaVision OMX V3 system (Applied Precision) equipped with a 100×/1.40 NA PlanApo oil immersion objective (Olympus) and Cascade II:512 EMCCD cameras (Photometrics) using 405-nm and 592-nm diode laser.

Mitotic abnormalities

To quantify the number of abnormal mitotic figures in MEF cells, mitotic cells were collected by shake-off, spun onto a glass-bottomed 96-black well plate (Greiner Bio-One) for 5 min at 400g, fixed in 4% formaldehyde, and stained with VectaShield containing DAPI.

Coimmunoprecipitation

Nuclei of wild-type and *Suv4-20h2*^{HA-Flag} ES cells were sequentially extracted with low-salt (fraction I) and high-salt (fraction II) immunoprecipitation buffer and incubated with 5 μg of either

Flag M2 antibody (Sigma) or Smc1 antibody (Bethyl Laboratories) coupled to ProteinA/G magnetic beads (Dynabeads, Invitrogen) overnight at 4°C on a rotating wheel. Bound proteins were separated on SDS–polyacrylamide gels and analyzed by Western blotting.

ChIP

ChIP experiments were performed as in Samoshkin et al. (2012) using the following antibodies: anti-H3K9me3 (Abcam, #ab8898), anti-H4K20me3 (Active Motif, #39180.3918; or Millipore, #07-463), anti-Rad21 (Abcam, #ab992), and anti-Smc3 (Abcam, #ab9263).

Automatic 3D quantification of FISH foci

Major satellite FISH experiments were performed according to Markaki et al. (2013). To determine quantitative information of FISH major satellite repeats from the acquired two-channel 3D confocal microscopy images, we developed a fully automatic 3D image analysis approach consisting of 3D cell nuclei segmentation and 3D foci segmentation.

Acknowledgments

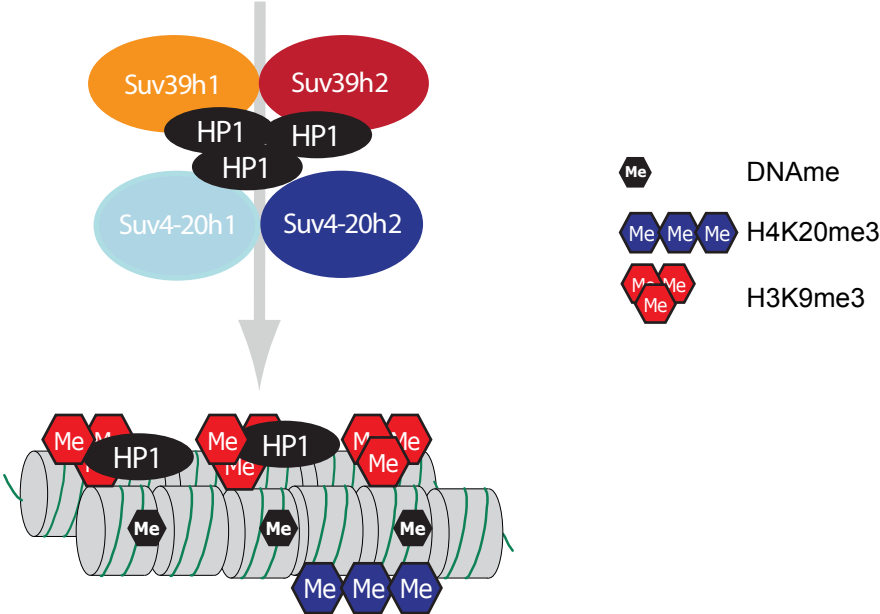
We thank Alexander Nuber for technical help. We are grateful to Thomas Cremer for sending major satellite FISH probes and protocols. We thank Jan-Michael Peters for sending antibodies, and Matthias Merkenschlager for sharing the Scc1 conditional knockout cell lines. We are indebted to Jerome Dejardin for discussions and sharing unpublished data. The work of G.S., K. Rippe, H.L., and K. Rohr was funded by the BMBF project EpiSys in the SysTec program. Additional support was from the Deutsche Forschungsgemeinschaft (SFB-TR5 and SFB684) to L.S., H.L., and G.S. H.v.M. was supported by the Fritz Thyssen Stiftung (AZ. 10.05.1.165).

References

- Bulut-Karslioglu A, Perrera V, Scaranaro M, de la Rosa-Velazquez IA, van de Nobelen S, Shukeir N, Popow J, Gerle B, Opravil S, Pagani M, et al. 2012. A transcription factor-based mechanism for mouse heterochromatin formation. *Nat Struct Mol Biol* **19**: 1023–1030.
- Cheutin T, McNairn AJ, Jenuwein T, Gilbert DM, Singh PB, Misteli T. 2003. Maintenance of stable heterochromatin domains by dynamic HP1 binding. *Science* **299**: 721–725.
- Degner SC, Verma-Gaur J, Wong TP, Bossen C, Iverson GM, Torkamani A, Vettermann C, Lin YC, Ju Z, Schulz D, et al. 2011. CCCTC-binding factor (CTCF) and cohesin influence the genomic architecture of the Igh locus and antisense transcription in pro-B cells. *Proc Natl Acad Sci* **108**: 9566–9571.
- Eckert CA, Gravidahl DJ, Megee PC. 2007. The enhancement of pericentromeric cohesin association by conserved kinetochore components promotes high-fidelity chromosome segregation and is sensitive to microtubule-based tension. *Genes Dev* **21**: 278–291.
- Fan Y, Nikitina T, Zhao J, Fleury TJ, Bhattacharyya R, Bouhassira EE, Stein A, Woodcock CL, Skoultschi AI. 2005. Histone H1 depletion in mammals alters global chromatin structure but causes specific changes in gene regulation. *Cell* **123**: 1199–1212.
- Fraga MF, Ballestar E, Villar-Garea A, Boix-Chornet M, Espada J, Schotta G, Bonaldi T, Haydon C, Roperio S, Petrie K, et al. 2005. Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nat Genet* **37**: 391–400.
- Gause M, Misulovin Z, Bilyeu A, Dorsett D. 2010. Dosage-sensitive regulation of cohesin chromosome binding and dynamics by Nipped-B, Pds5, and Wapl. *Mol Cell Biol* **30**: 4940–4951.
- Gilbert N, Thomson I, Boyle S, Allan J, Ramsahoye B, Bickmore WA. 2007. DNA methylation affects nuclear organization, histone modifications, and linker histone binding but not chromatin compaction. *J Cell Biol* **177**: 401–411.
- Goodarzi AA, Noon AT, Deckbar D, Ziv Y, Shiloh Y, Lobrich M, Jeggo PA. 2008. ATM signaling facilitates repair of DNA double-strand breaks associated with heterochromatin. *Mol Cell* **31**: 167–177.
- Guenatri M, Bailly D, Maison C, Almouzni G. 2004. Mouse centric and pericentric satellite repeats form distinct functional heterochromatin. *J Cell Biol* **166**: 493–505.
- Hadjur S, Williams LM, Ryan NK, Cobb BS, Sexton T, Fraser P, Fisher AG, Merkenschlager M. 2009. Cohesins form chromosomal *cis*-interactions at the developmentally regulated IFNG locus. *Nature* **460**: 410–413.
- Hahn M, Dambacher S, Schotta G. 2010. Heterochromatin dysregulation in human diseases. *J Appl Physiol* **109**: 232–242.
- Hauf S, Waizenegger IC, Peters JM. 2001. Cohesin cleavage by separase required for anaphase and cytokinesis in human cells. *Science* **293**: 1320–1323.
- Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, Ebmeier CC, Goossens J, Rahl PB, Levine SS, et al. 2010. Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**: 430–435.
- Koch B, Kueng S, Ruckenbauer C, Wendt KS, Peters JM. 2008. The Suv39h–HP1 histone methylation pathway is dispensable for enrichment and protection of cohesin at centromeres in mammalian cells. *Chromosoma* **117**: 199–210.
- Kueng S, Hegemann B, Peters BH, Lipp JJ, Schleiffer A, Mechtler K, Peters JM. 2006. Wapl controls the dynamic association of cohesin with chromatin. *Cell* **127**: 955–967.
- Lengronne A, Katou Y, Mori S, Yokobayashi S, Kelly GP, Itoh T, Watanabe Y, Shirahige K, Uhlmann F. 2004. Cohesin relocation from sites of chromosomal loading to places of convergent transcription. *Nature* **430**: 573–578.
- Markaki Y, Smeets D, Cremer M, Schermelleh L. 2013. Fluorescence in situ hybridization applications for super-resolution 3D structured illumination microscopy. *Methods Mol Biol* **950**: 43–64.
- Meshorer E, Yellajoshula D, George E, Scambler PJ, Brown DT, Misteli T. 2006. Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev Cell* **10**: 105–116.
- Muller KP, Erdel F, Caudron-Herger M, Marth C, Fodor BD, Richter M, Scaranaro M, Beaudouin J, Wachsmuth M, Rippe K. 2009. Multiscale analysis of dynamics and interactions of heterochromatin protein 1 by fluorescence fluctuation microscopy. *Biophys J* **97**: 2876–2885.
- Nasmyth K. 2011. Cohesin: A catenase with separate entry and exit gates? *Nat Cell Biol* **13**: 1170–1177.
- Nicetto D, Hahn M, Jung J, Schneider TD, Straub T, David R, Schotta G, Rupp RA. 2013. Suv4-20h histone methyltransferases promote neuroectodermal differentiation by silencing the pluripotency-associated oct-25 gene. *PLoS Genet* **9**: e1003188.
- Nonaka N, Kitajima T, Yokobayashi S, Xiao G, Yamamoto M, Grewal SI, Watanabe Y. 2002. Recruitment of cohesin to heterochromatic regions by Swi6/HP1 in fission yeast. *Nat Cell Biol* **4**: 89–93.

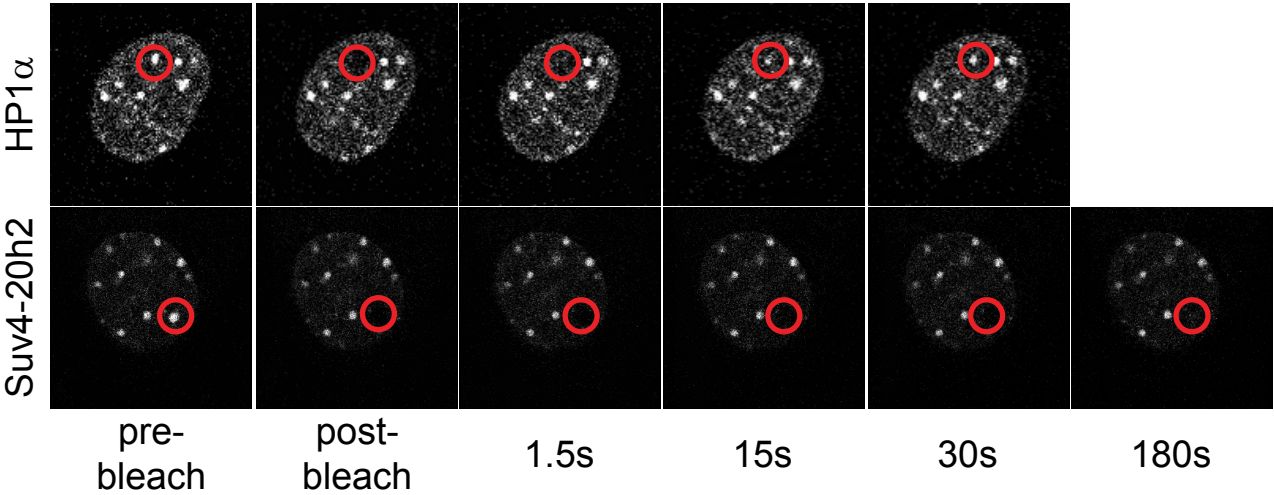
- Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, Gregson HC, Jarmuz A, Canzonetta C, Webster Z, Nesterova T, et al. 2008. Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* **132**: 422–433.
- Peters AH, O'Carroll D, Scherthan H, Mechtler K, Sauer S, Schofer C, Weipoltshammer K, Pagani M, Lachner M, Kohlmaier A, et al. 2001. Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* **107**: 323–337.
- Rubio ED, Reiss DJ, Welcsh PL, Distèche CM, Filippova GN, Baliga NS, Aebersold R, Ranish JA, Krumm A. 2008. CTCF physically links cohesin to chromatin. *Proc Natl Acad Sci* **105**: 8309–8314.
- Salic A, Waters JC, Mitchison TJ. 2004. Vertebrate shugoshin links sister centromere cohesion and kinetochore microtubule stability in mitosis. *Cell* **118**: 567–578.
- Samoshkin A, Dulev S, Loukinov D, Rosenfeld JA, Strunnikov AV. 2012. Condensin dysfunction in human cells induces nonrandom chromosomal breaks in anaphase, with distinct patterns for both unique and repeated genomic regions. *Chromosoma* **121**: 191–199.
- Schermelleh L, Carlton PM, Haase S, Shao L, Winoto L, Kner P, Burke B, Cardoso MC, Agard DA, Gustafsson MG, et al. 2008. Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy. *Science* **320**: 1332–1336.
- Schneider AC, Heukamp LC, Rogenhofer S, Fechner G, Bastian PJ, von Ruecker A, Muller SC, Ellinger J. 2011. Global histone H4K20 trimethylation predicts cancer-specific survival in patients with muscle-invasive bladder cancer. *BJU Int* **108**: E290–E296.
- Schotta G, Lachner M, Sarma K, Ebert A, Sengupta R, Reuter G, Reinberg D, Jenuwein T. 2004. A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev* **18**: 1251–1262.
- Schotta G, Sengupta R, Kubicek S, Malin S, Kauer M, Callen E, Celeste A, Pagani M, Opravil S, De La Rosa-Velazquez IA, et al. 2008. A chromatin-wide transition to H4K20 monomethylation impairs genome integrity and programmed DNA rearrangements in the mouse. *Genes Dev* **22**: 2048–2061.
- Seitan VC, Merckenschlager M. 2012. Cohesin and chromatin organisation. *Curr Opin Genet Dev* **22**: 93–100.
- Souza PP, Volkel P, Trinel D, Vandamme J, Rosnoblet C, Heliot L, Angrand PO. 2009. The histone methyltransferase SUV420H2 and heterochromatin proteins HP1 interact but show different dynamic behaviours. *BMC Cell Biol* **10**: 41.
- Tang Z, Sun Y, Harley SE, Zou H, Yu H. 2004. Human Bub1 protects centromeric sister-chromatid cohesion through Shugoshin during mitosis. *Proc Natl Acad Sci* **101**: 18012–18017.
- Thompson SL, Bakhoum SE, Compton DA. 2010. Mechanisms of chromosomal instability. *Curr Biol* **20**: R285–R295.
- Ting DT, Lipson D, Paul S, Brannigan BW, Akhavanfard S, Coffman EJ, Contino G, Deshpande V, Iafrate AJ, Letovsky S, et al. 2011. Aberrant overexpression of satellite repeats in pancreatic and other epithelial cancers. *Science* **331**: 593–596.
- Van Den Broeck A, Brambilla E, Moro-Sibilot D, Lantuejoul S, Brambilla C, Eymin B, Khochbin S, Gazzeri S. 2008. Loss of histone H4K20 trimethylation occurs in preneoplasia and influences prognosis of non-small cell lung cancer. *Clin Cancer Res* **14**: 7237–7245.
- Vermeulen M, Eberl HC, Matarese F, Marks H, Denissov S, Butter F, Lee KK, Olsen JV, Hyman AA, Stunnenberg HG, et al. 2010. Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* **142**: 967–980.
- Watrin E, Schleiffer A, Tanaka K, Eisenhaber F, Nasmyth K, Peters JM. 2006. Human Scc4 is required for cohesin binding to chromatin, sister-chromatid cohesion, and mitotic progression. *Curr Biol* **16**: 863–874.
- Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, Tsutsumi S, Nagae G, Ishihara K, Mishiro T, et al. 2008. Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* **451**: 796–801.
- Zhu Q, Pao GM, Huynh AM, Suh H, Tonnu N, Nederlof PM, Gage FH, Verma IM. 2011. BRCA1 tumour suppression occurs via heterochromatin-mediated silencing. *Nature* **477**: 179–184.

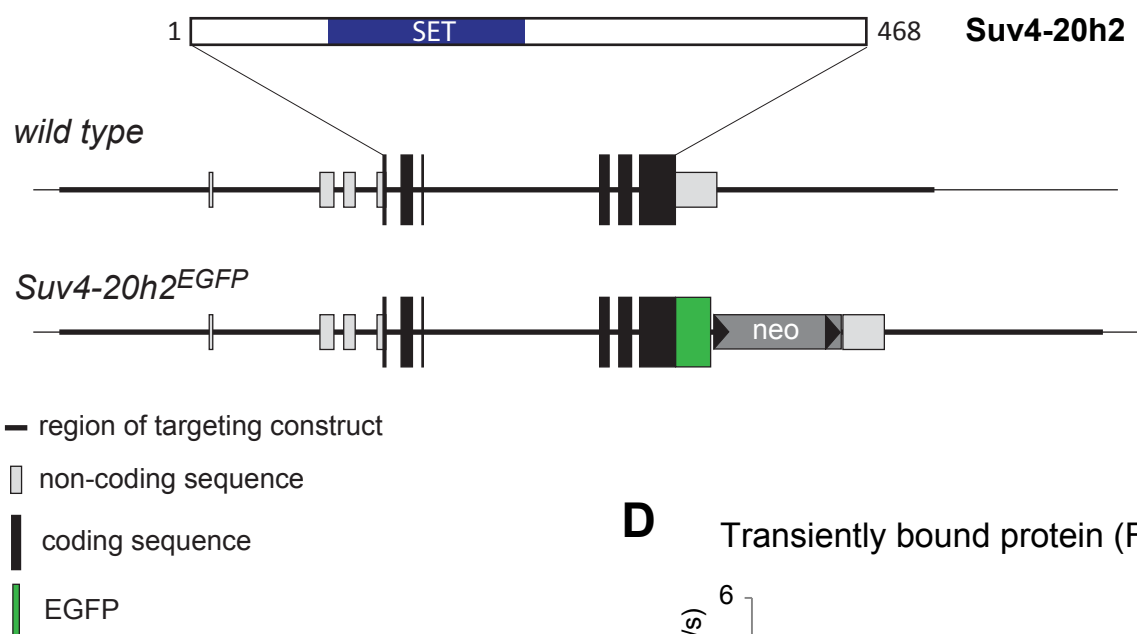
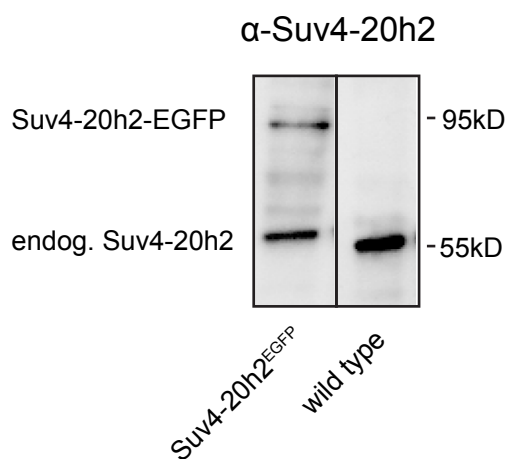
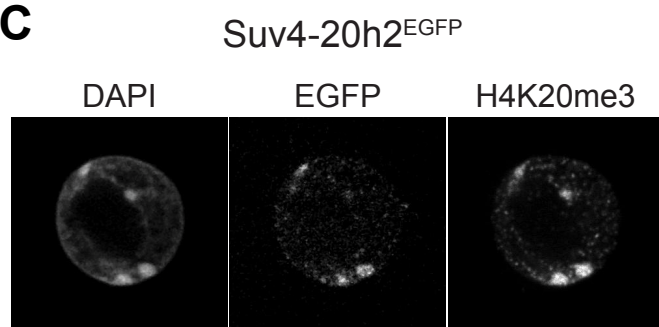
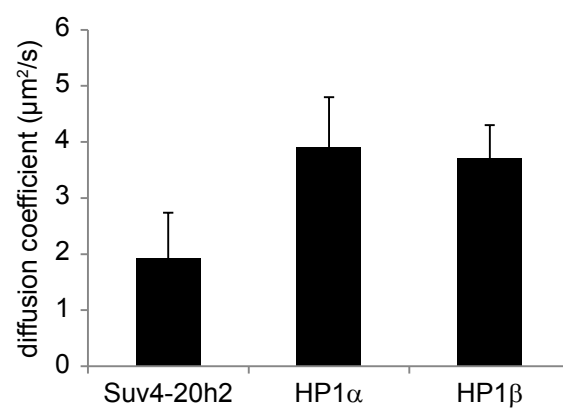
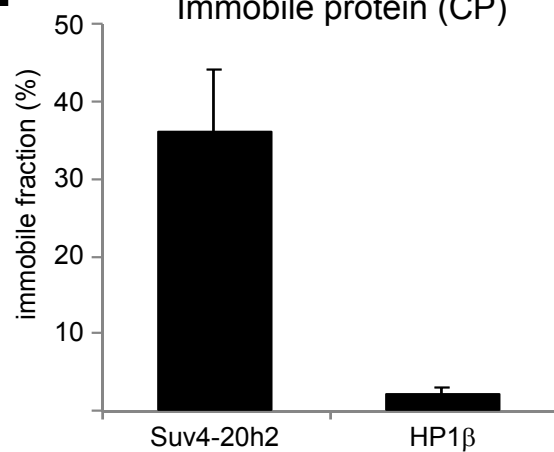
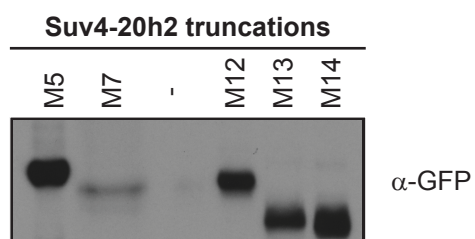
A

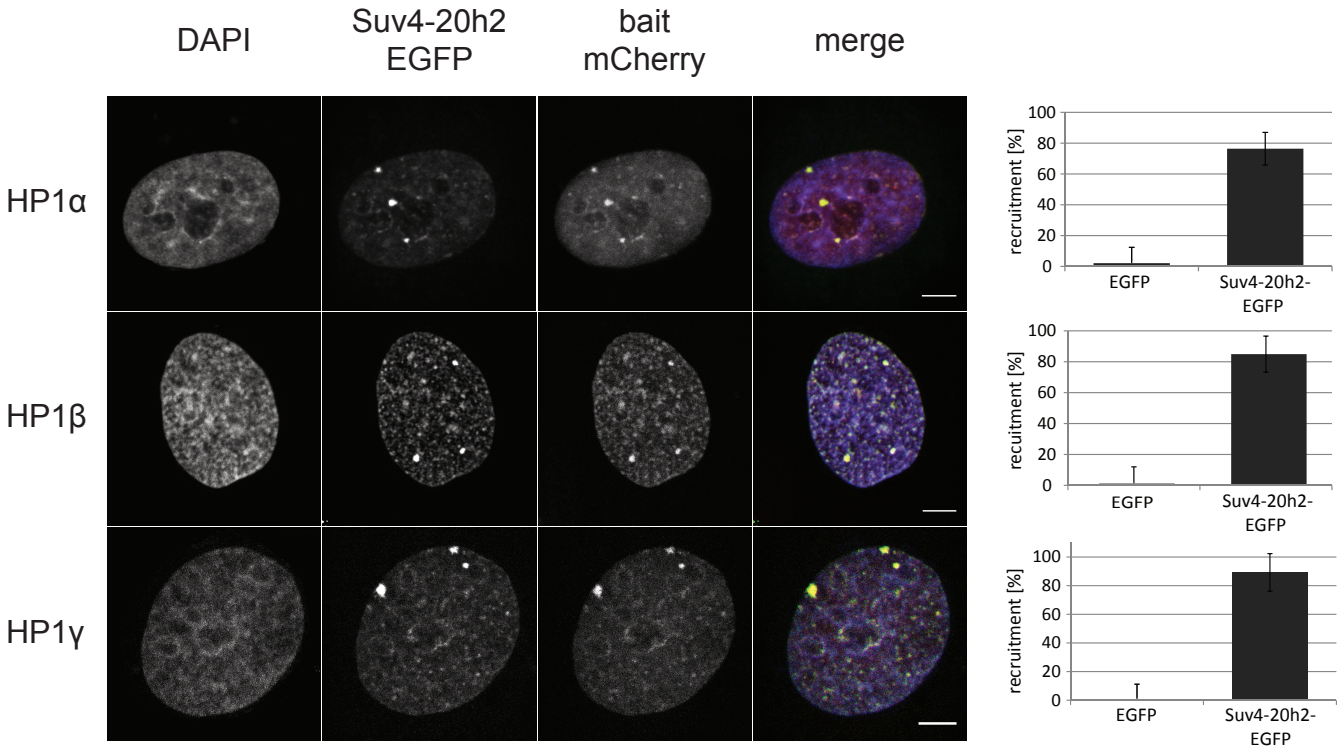


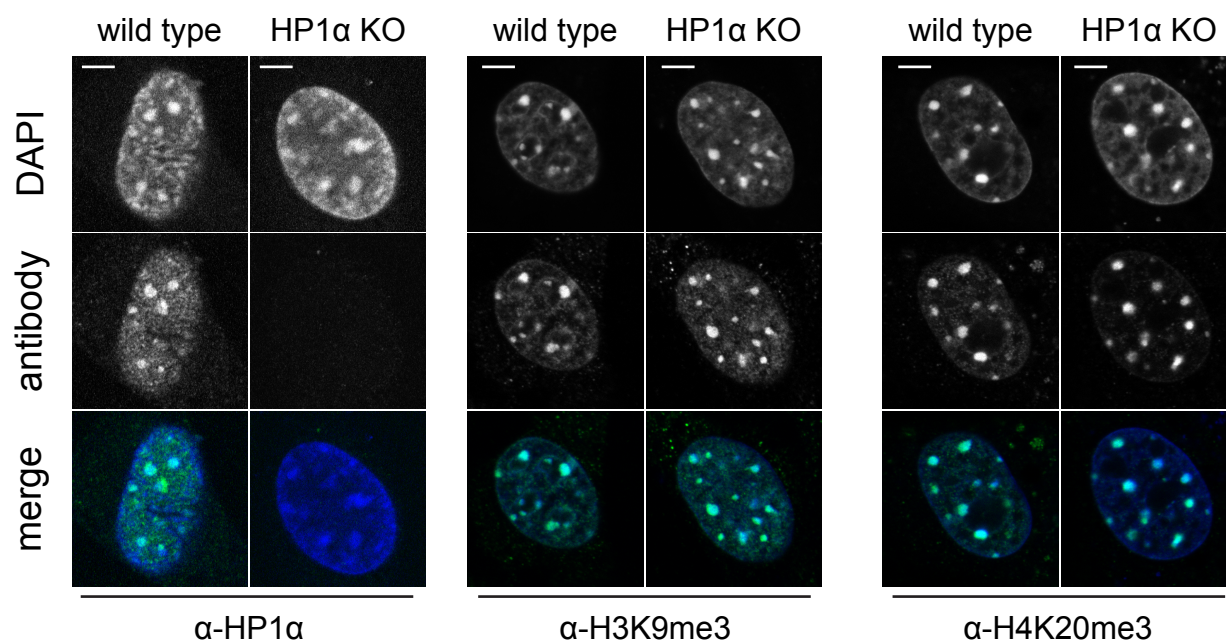
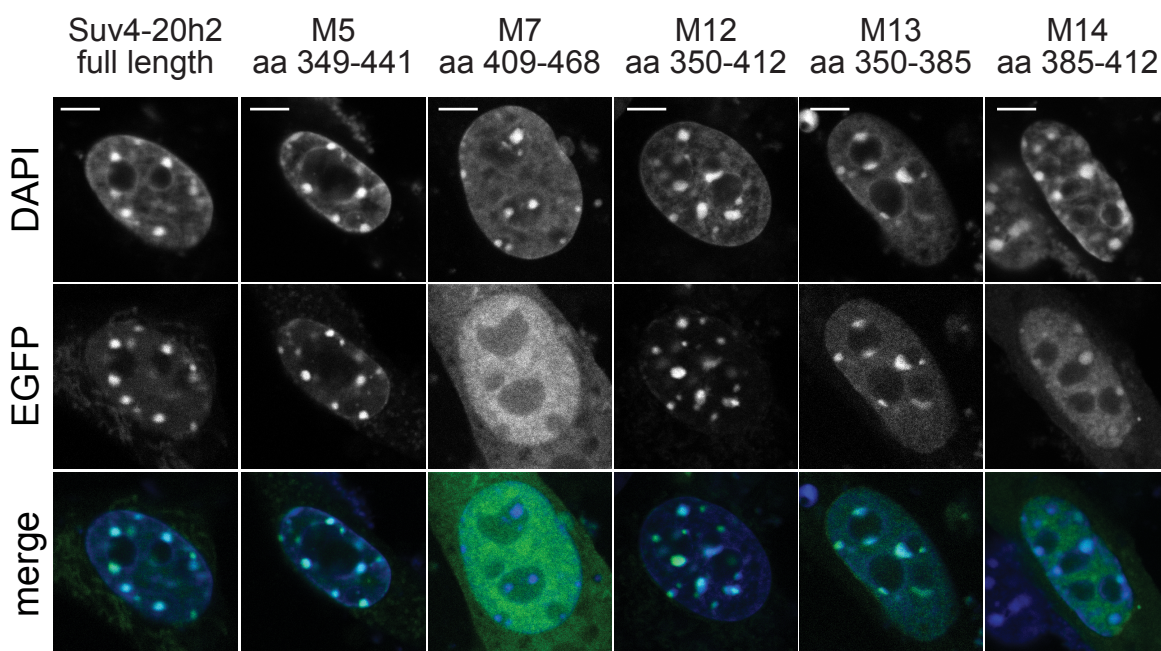
B

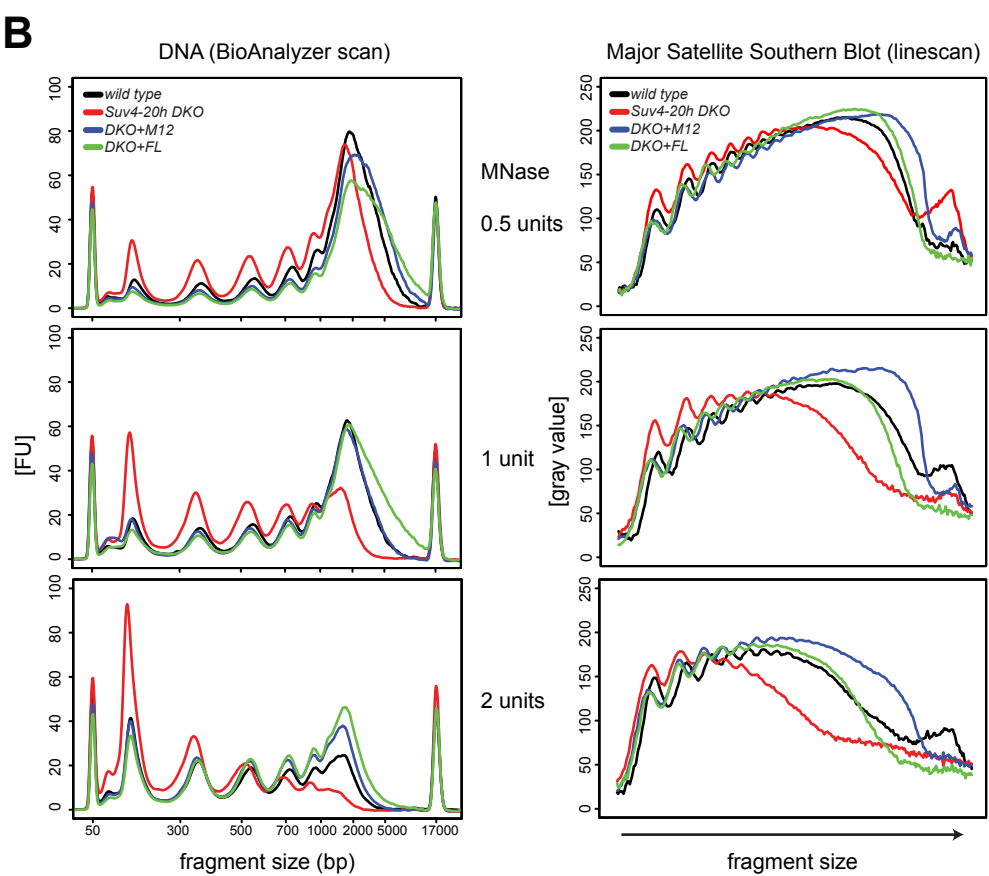
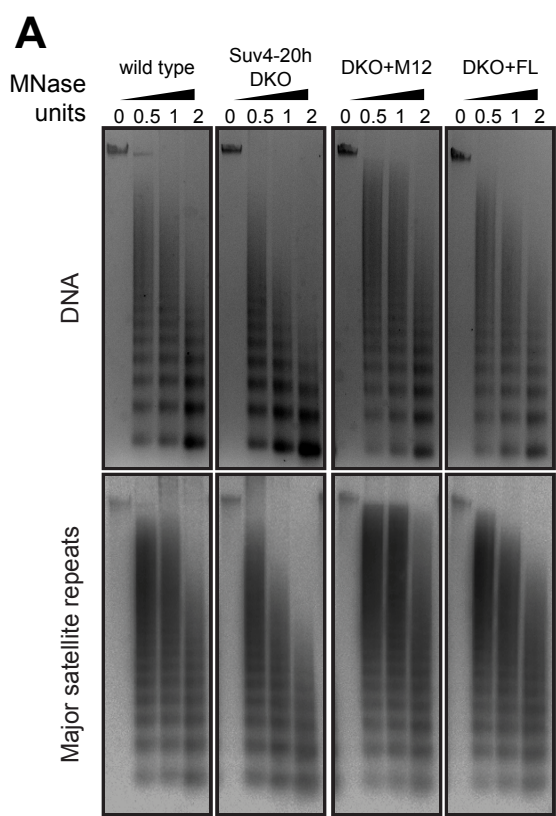
FRAP analysis at heterochromatin

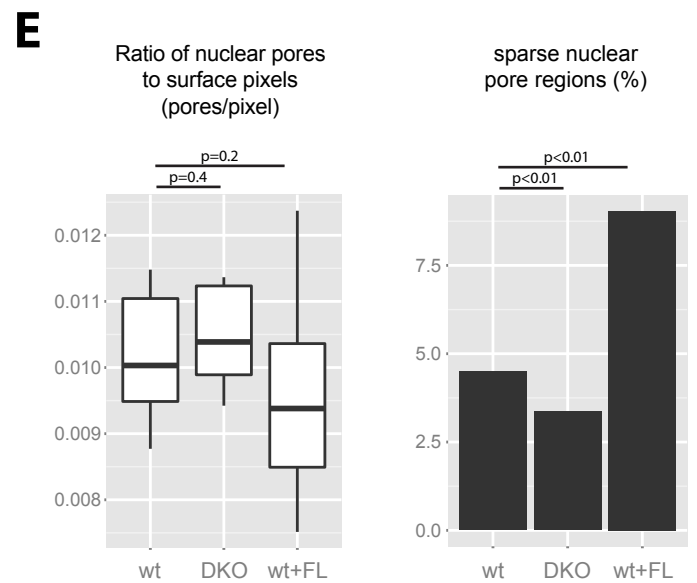
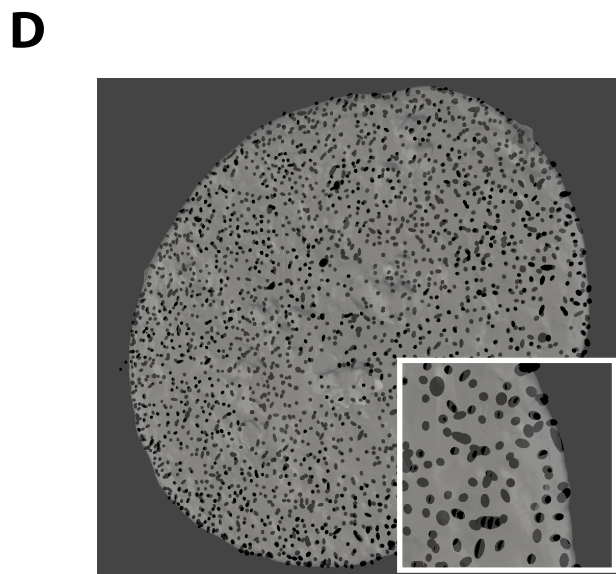
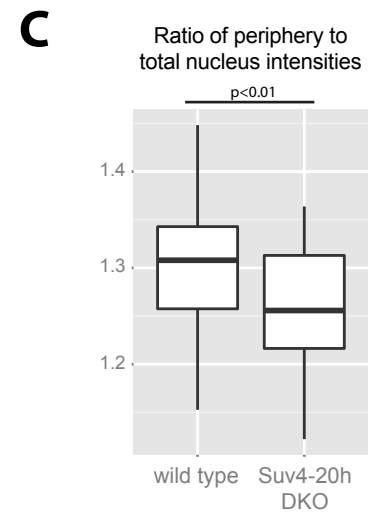
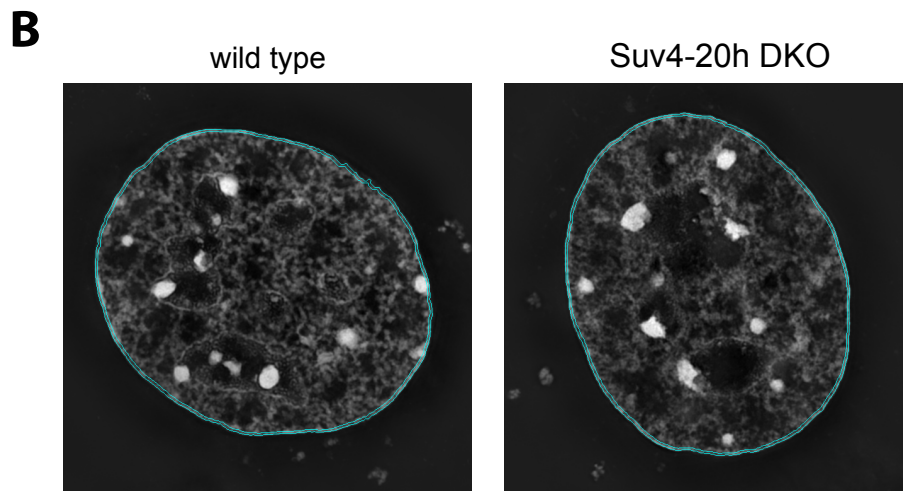
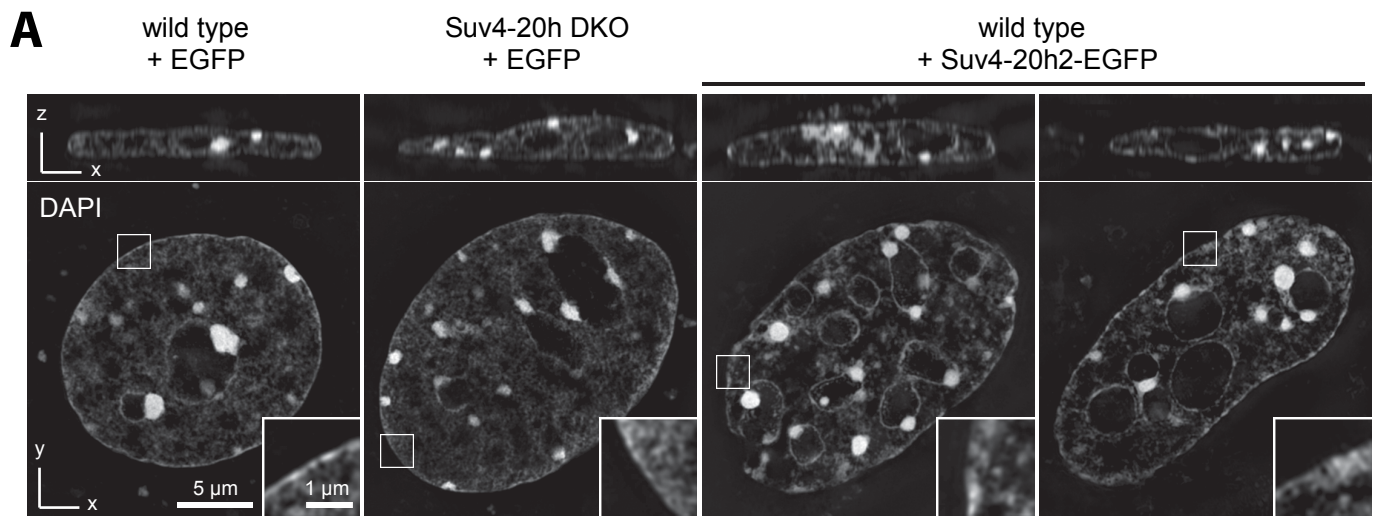


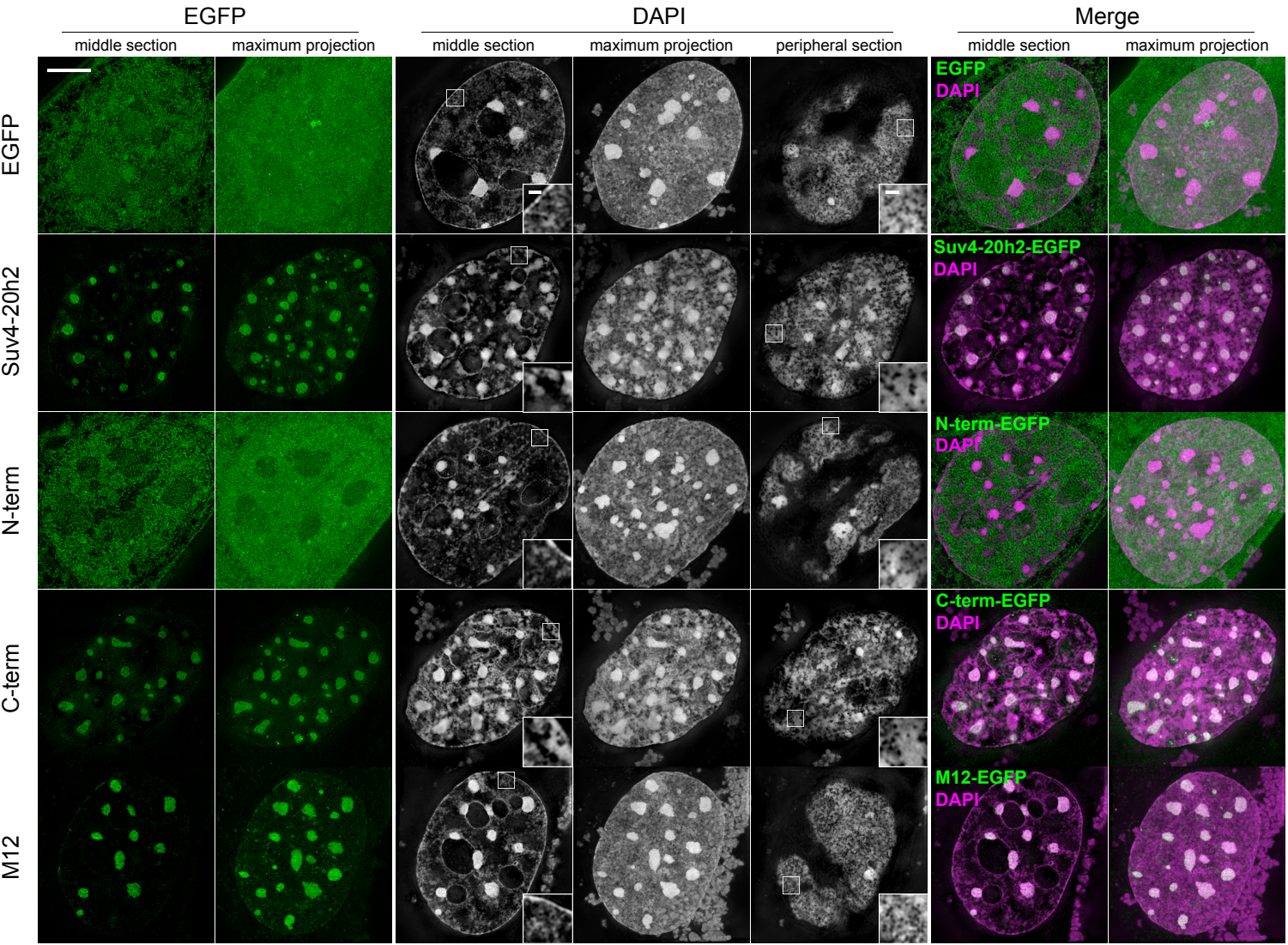
A**Suv4-20h2 knock-in strategy****B****C****D****Transiently bound protein (FCS)****E****Immobile protein (CP)****F**

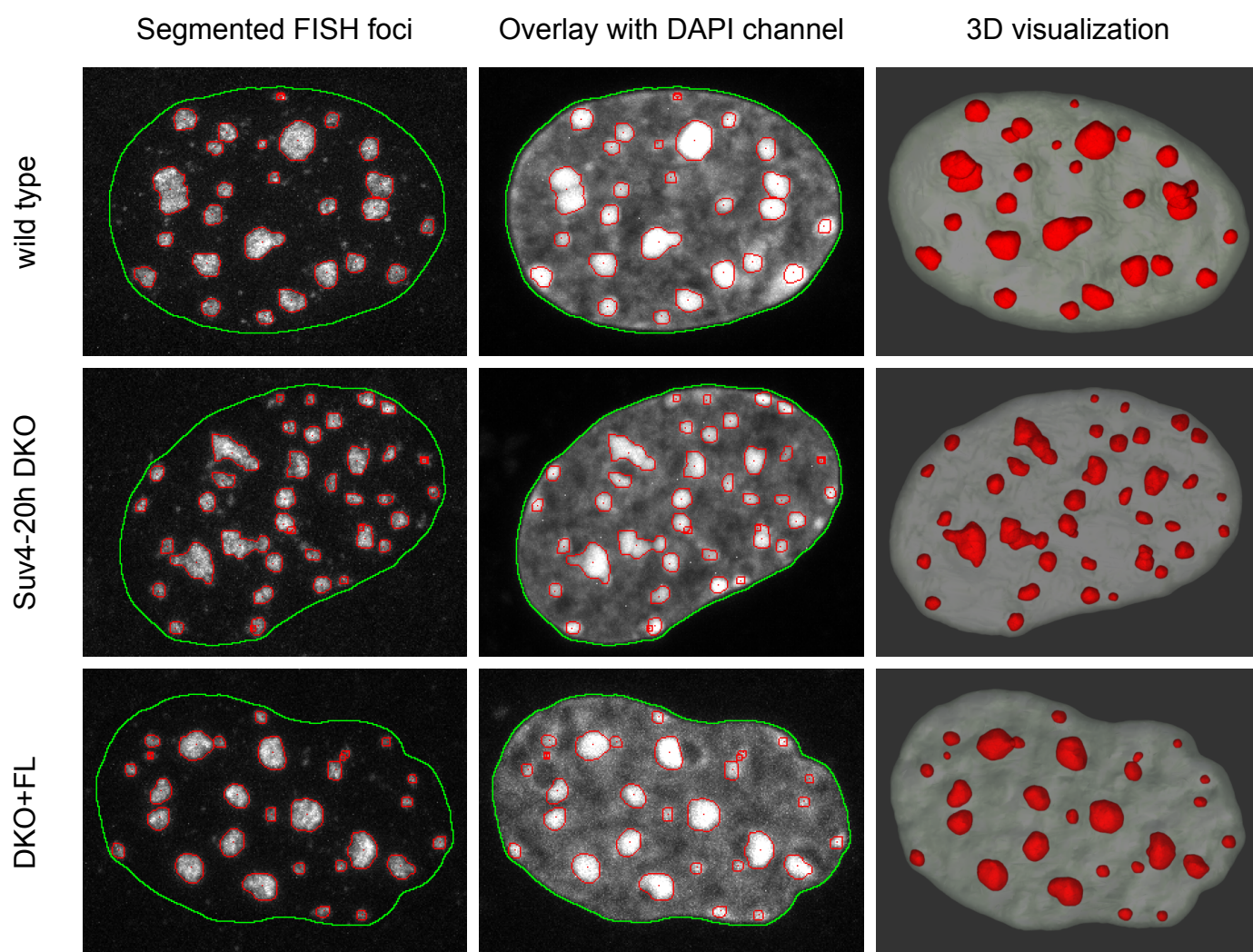


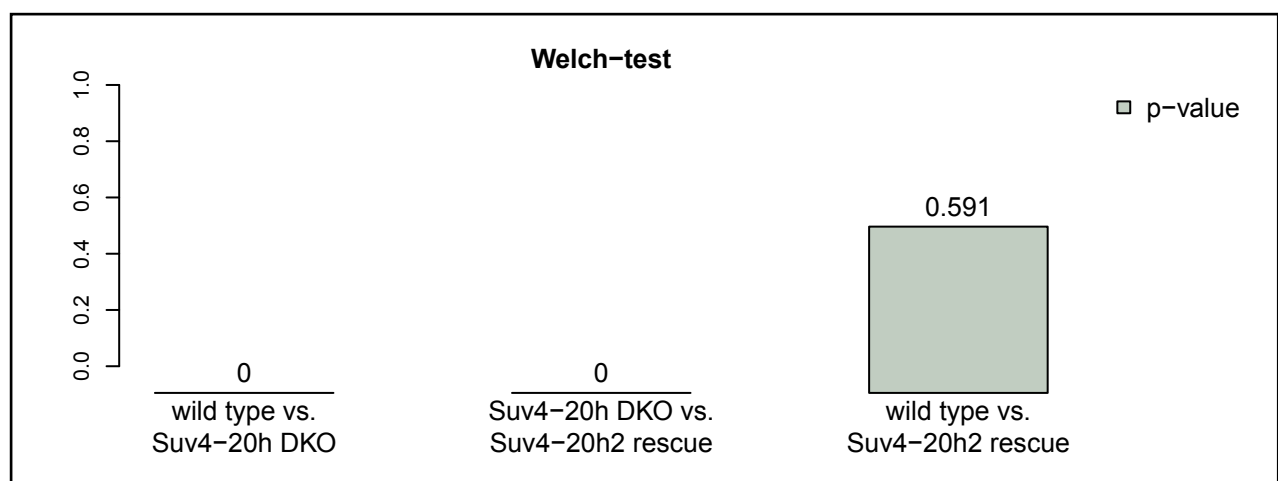
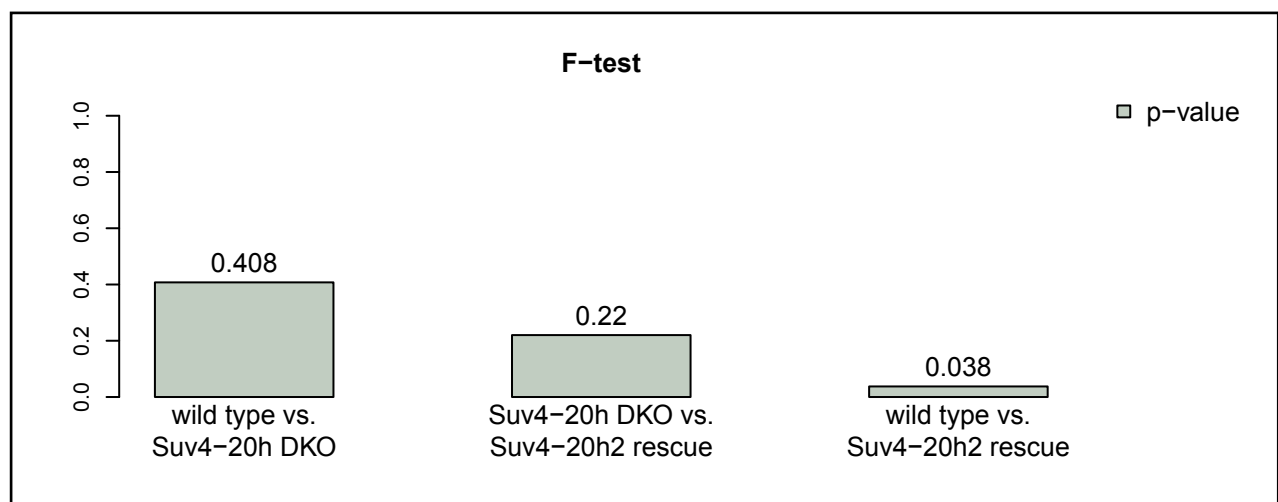
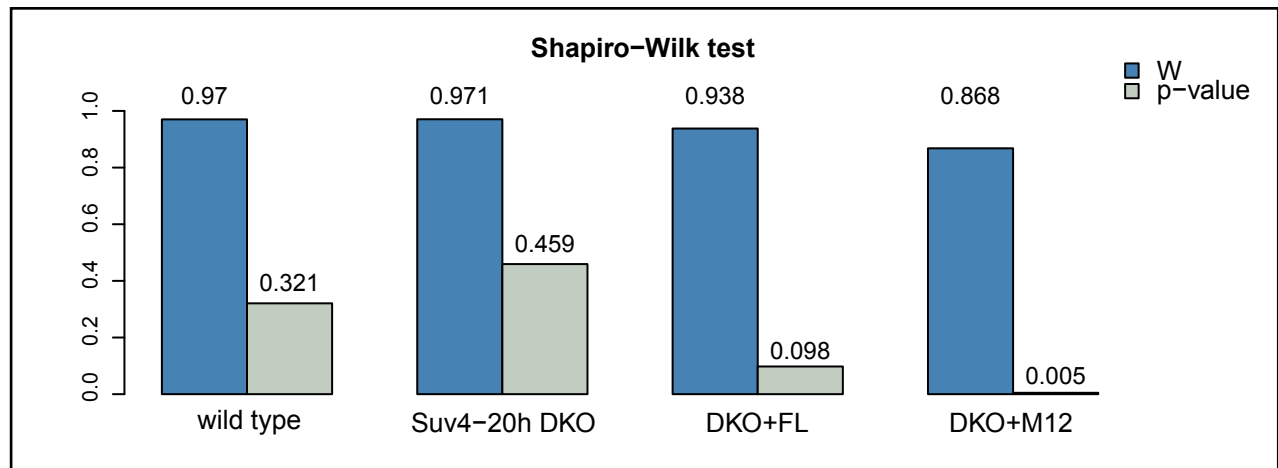
A**B****HP1 α ko fibroblasts**

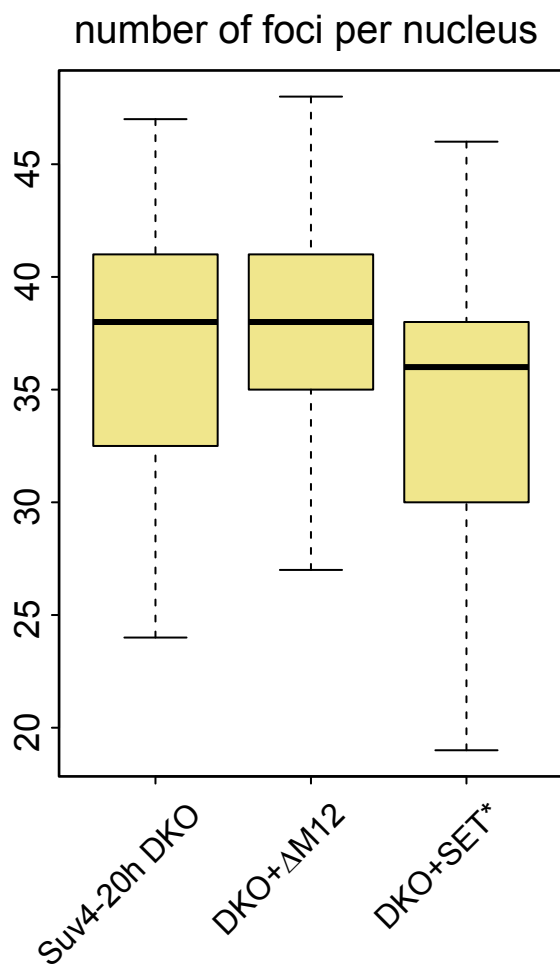












Shapiro-Wilk test

DKO: W=0,9519
p-value=0,2205

DKO+ΔM12: W=0,9705
p-value=0,6575

DKO+SET*: W=0,9591
p-value=0,3963

F-test

DKO
vs. DKO+ΔM12: p-value=0,6402

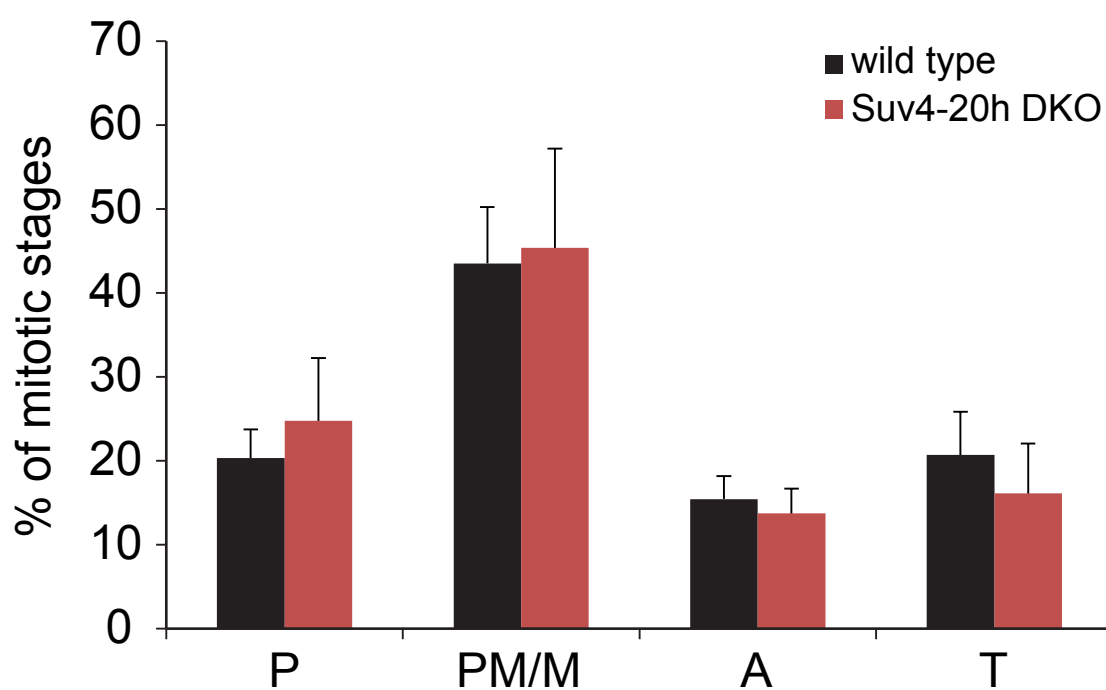
DKO
vs. DKO+SET*: p-value=0,8246

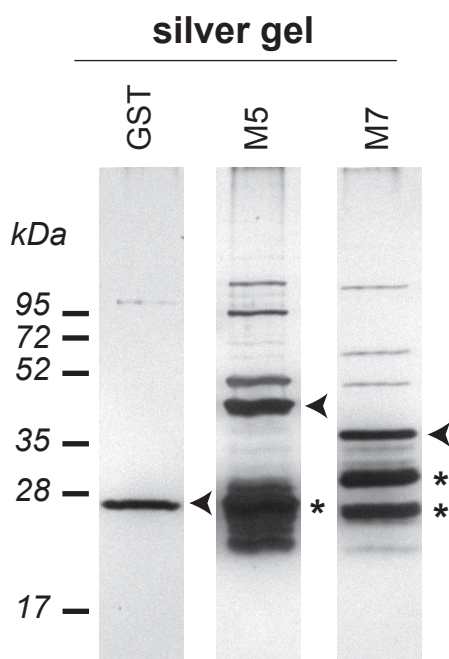
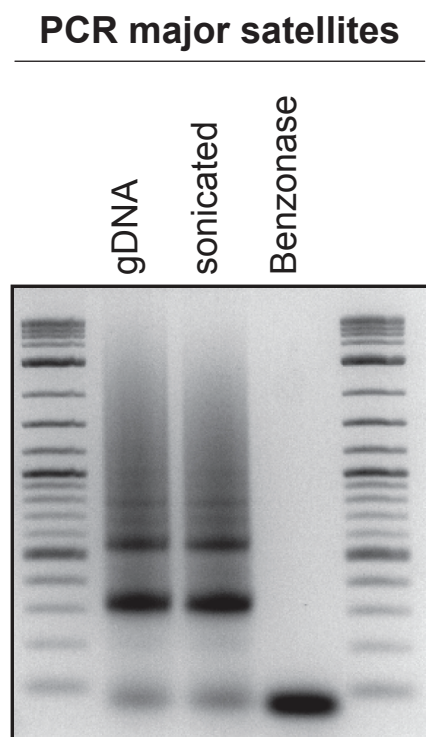
Welch-test

DKO
vs. DKO+ΔM12: p-value=0,4517

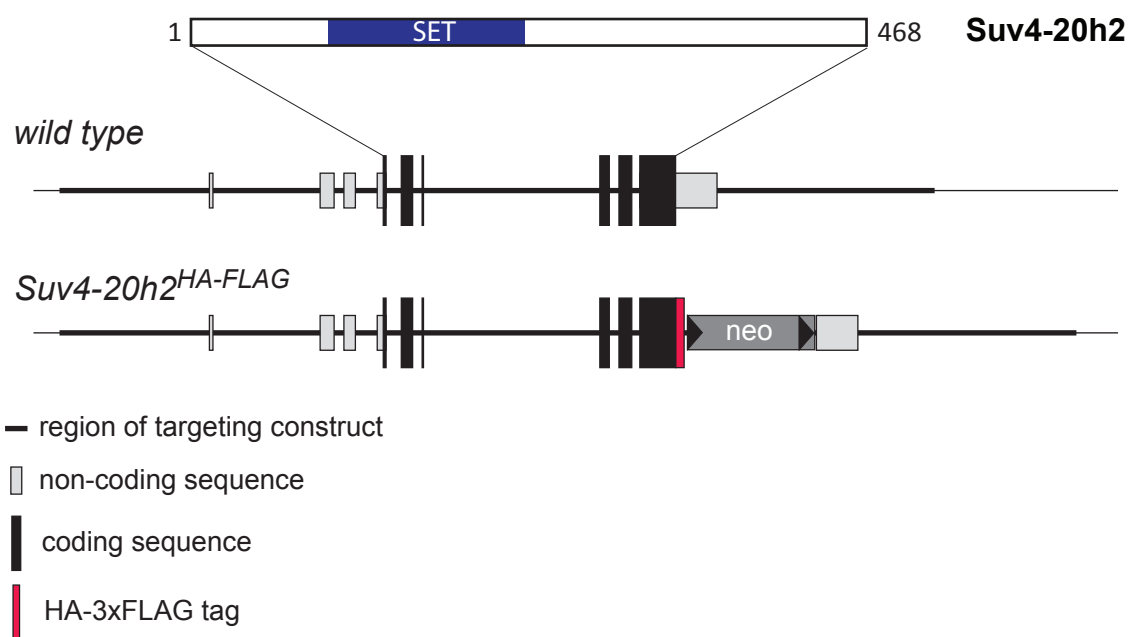
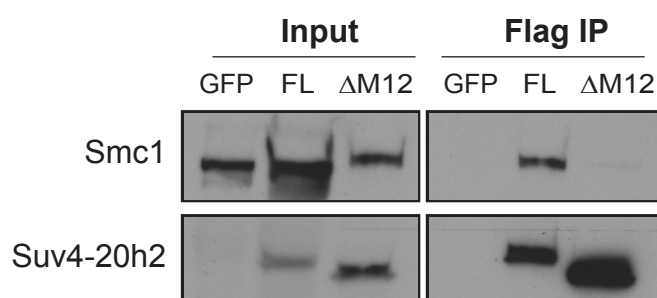
DKO
vs. DKO+SET*: p-value=0,1573

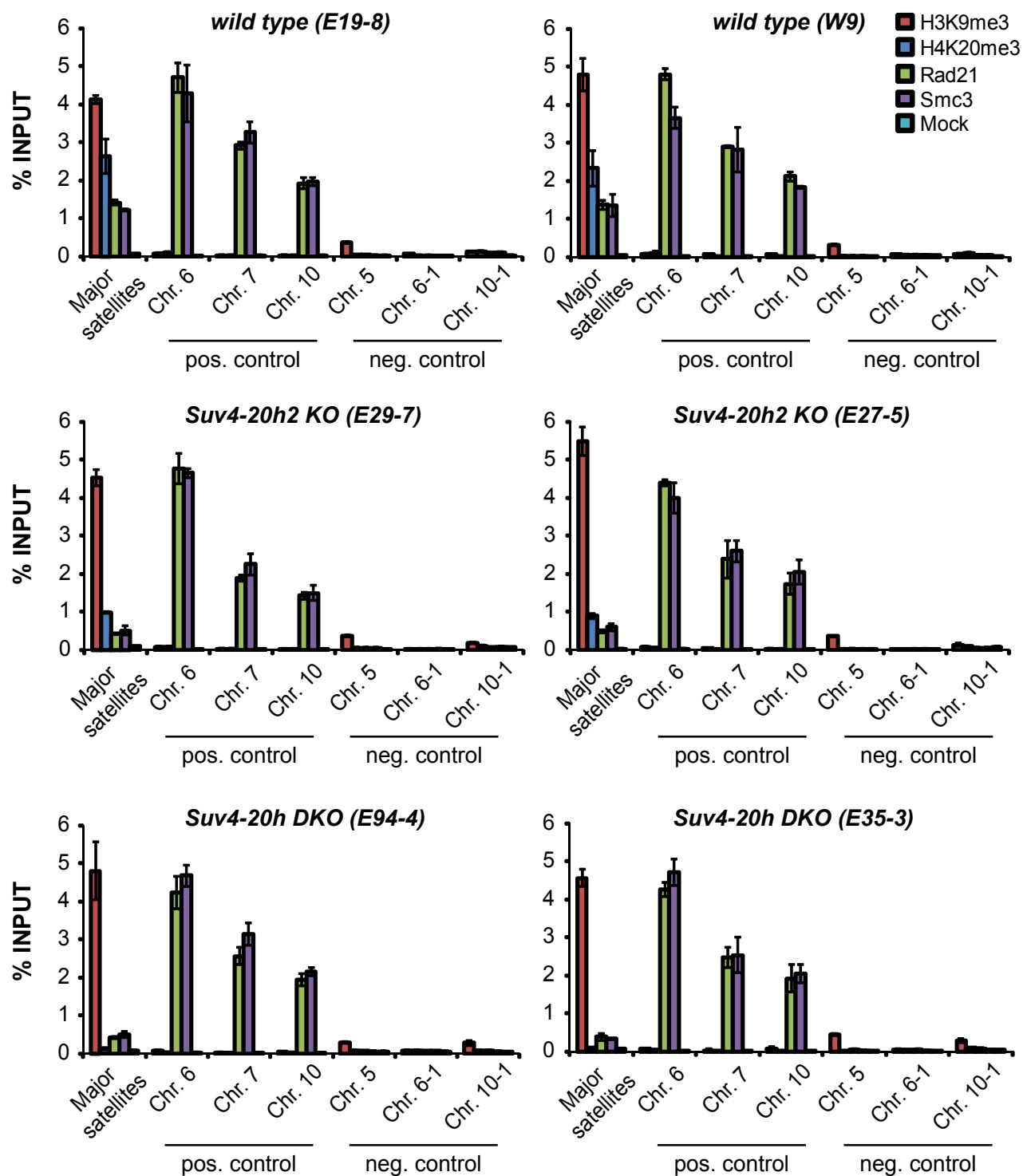
mitotic stages



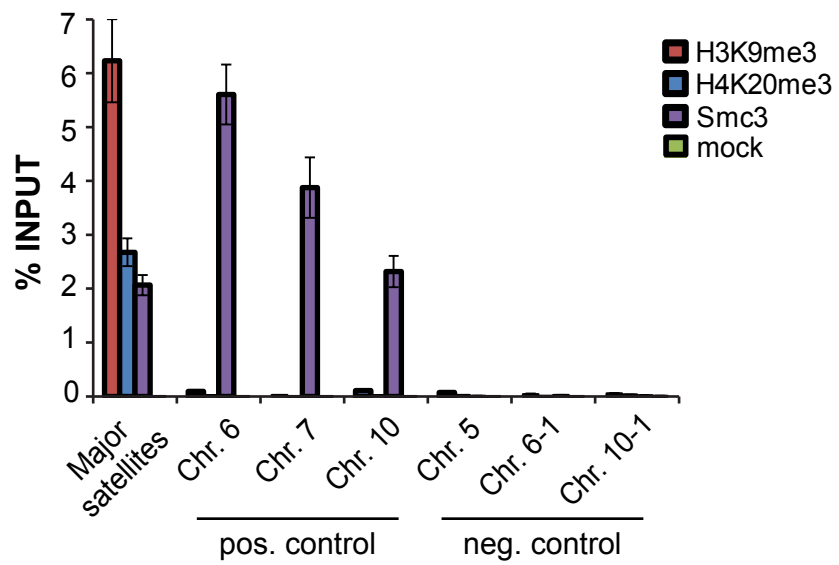
A**B****C**

Suv4-20h2 knock-in strategy

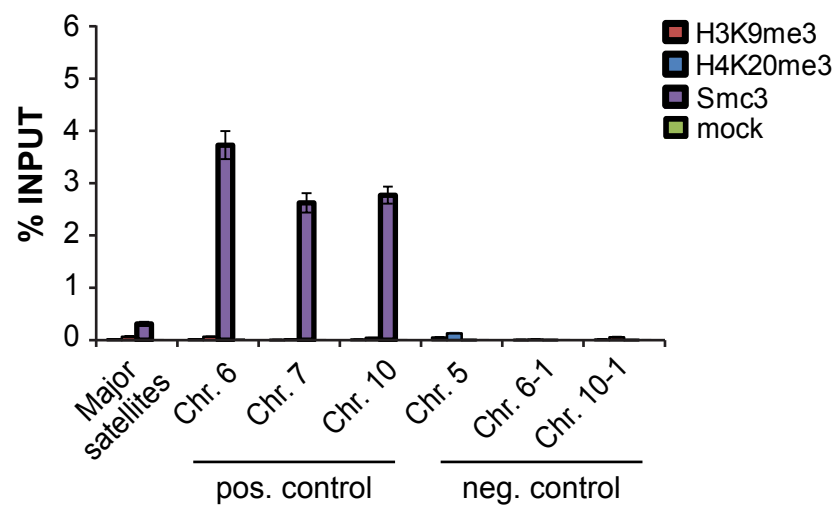
**D**



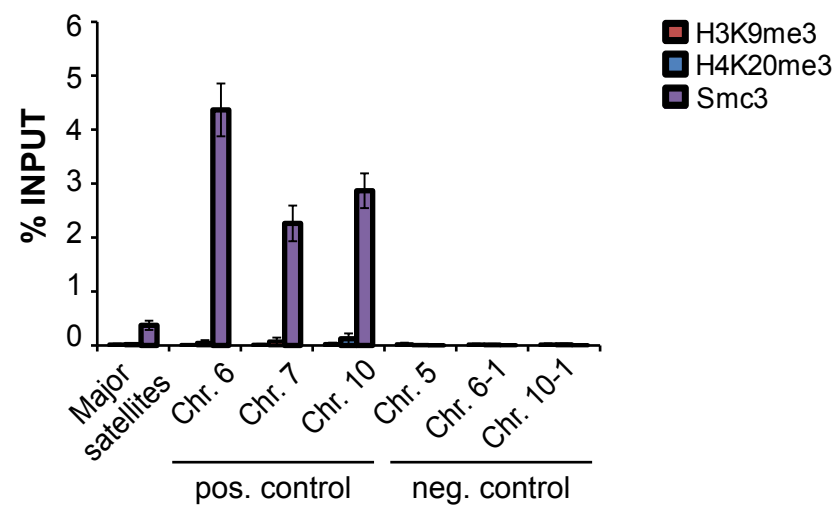
**wild type
(W9)**

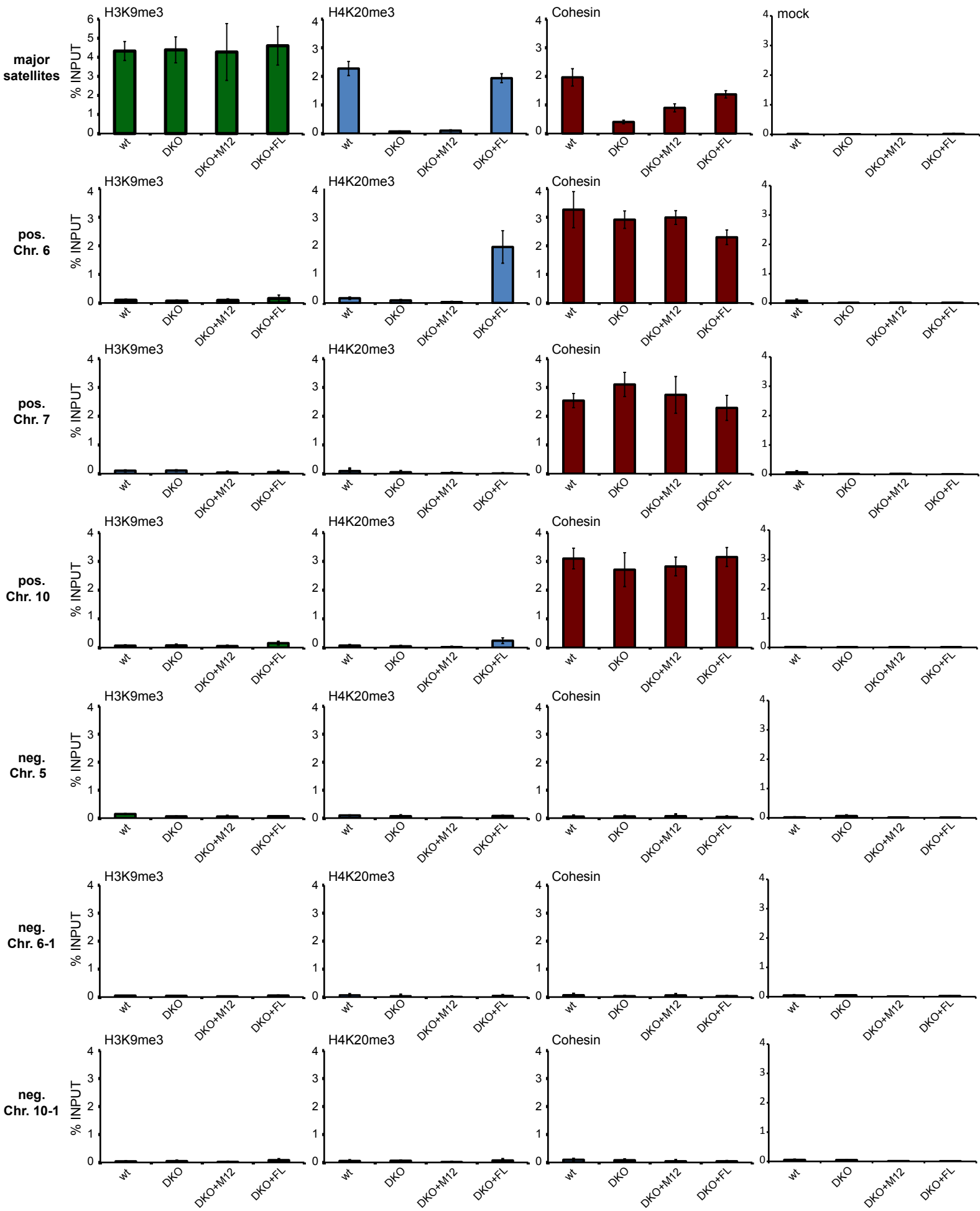


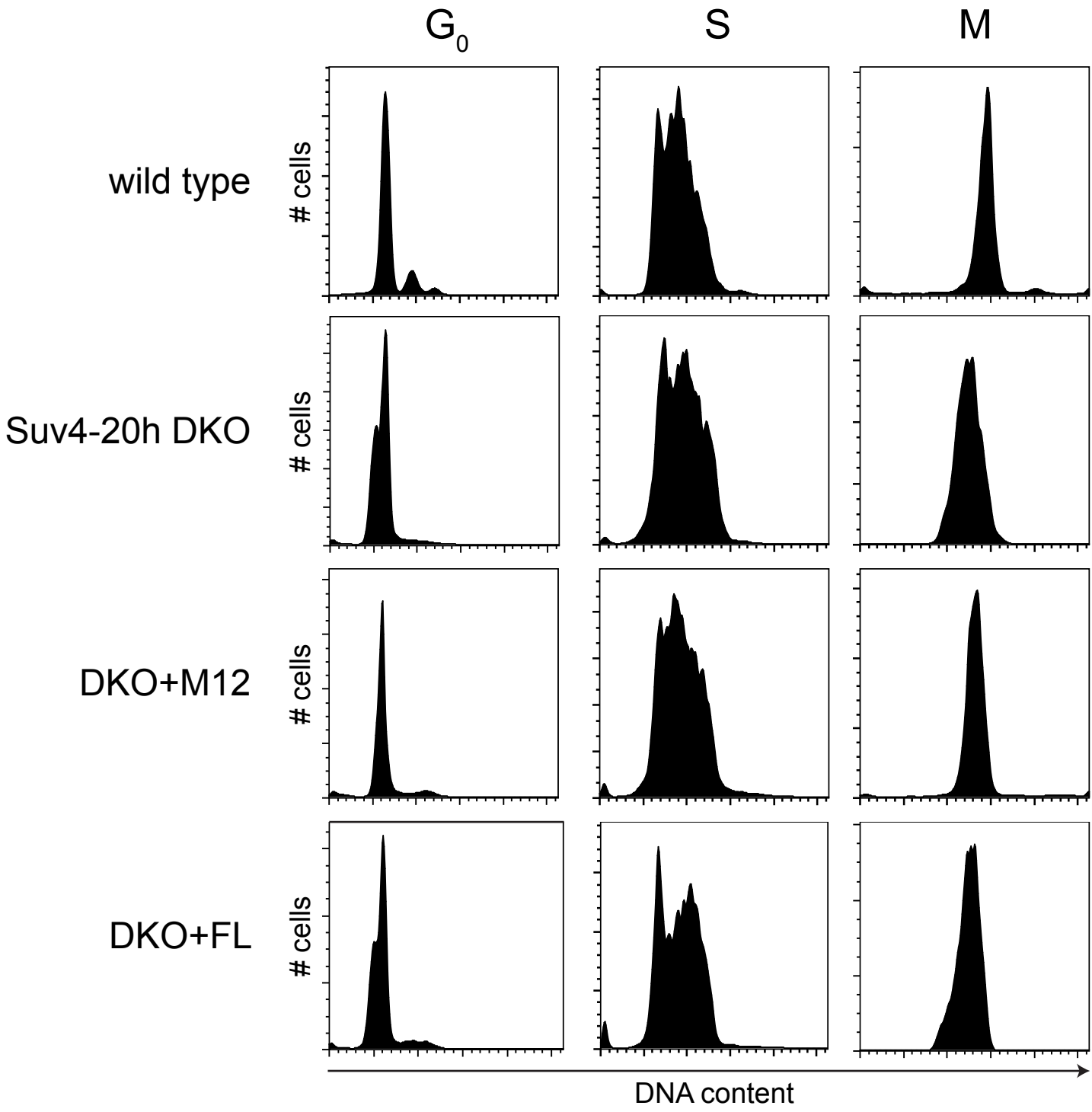
**Suv39h DKO
(E124-1)**

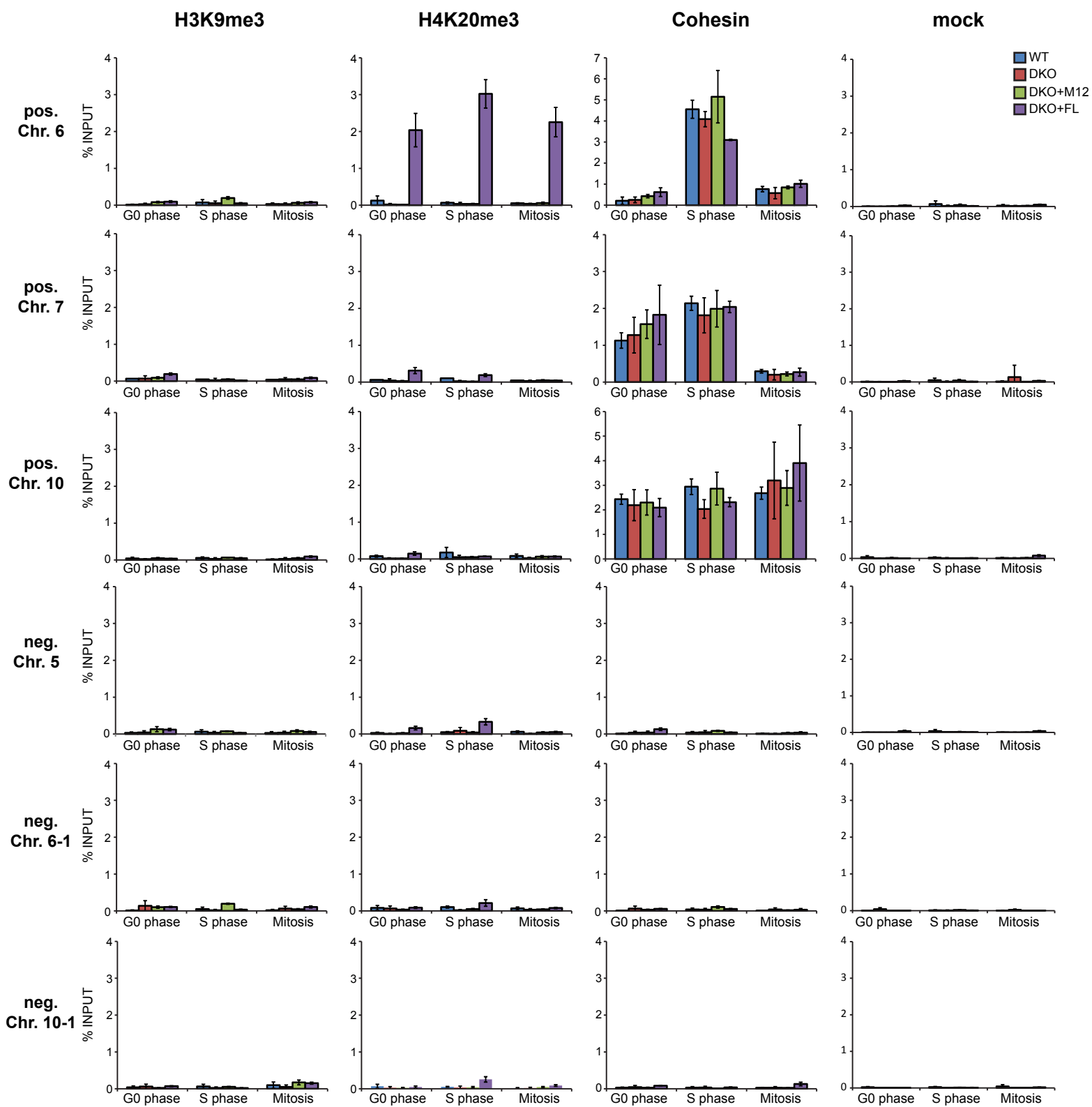


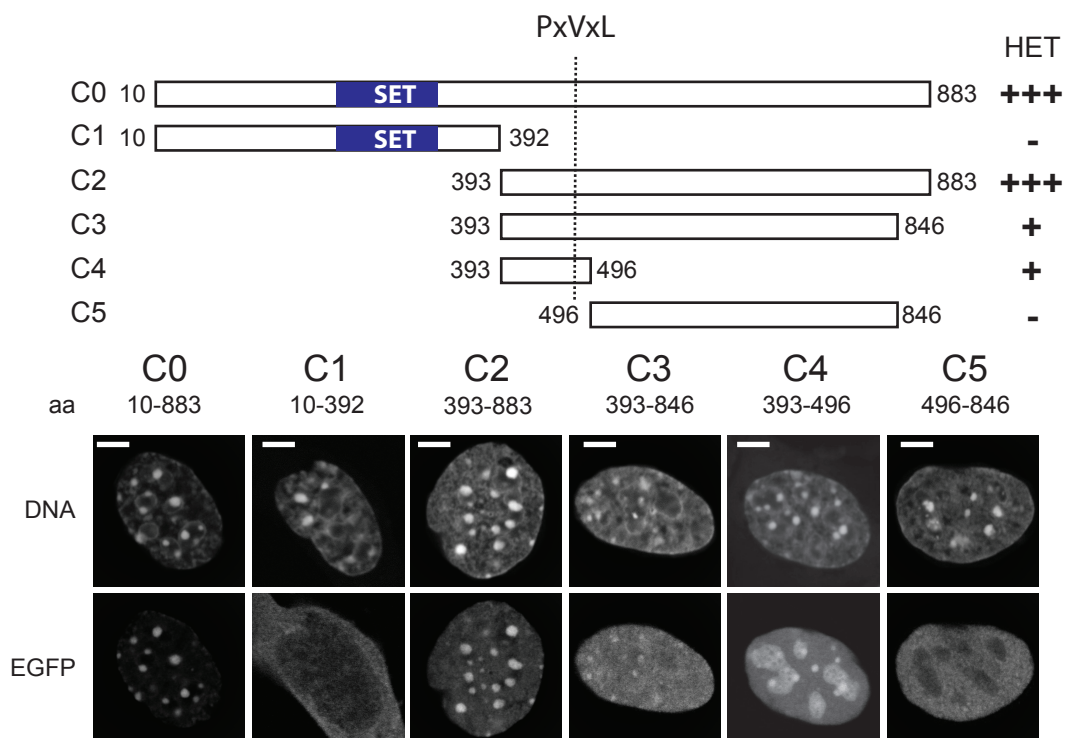
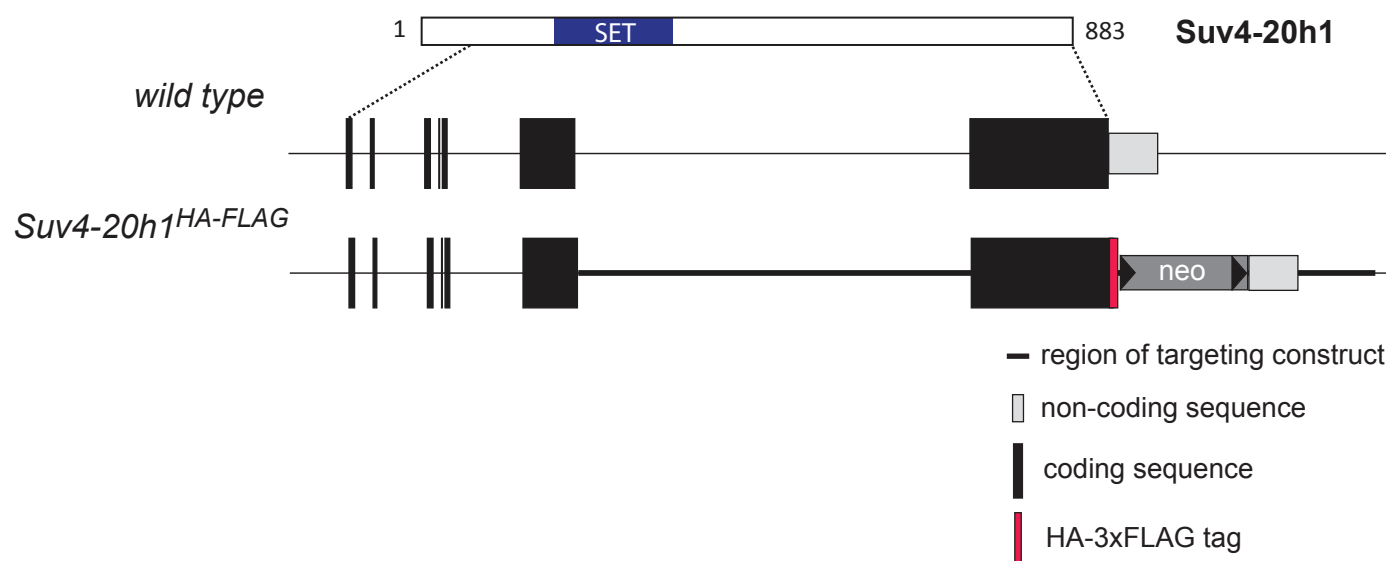
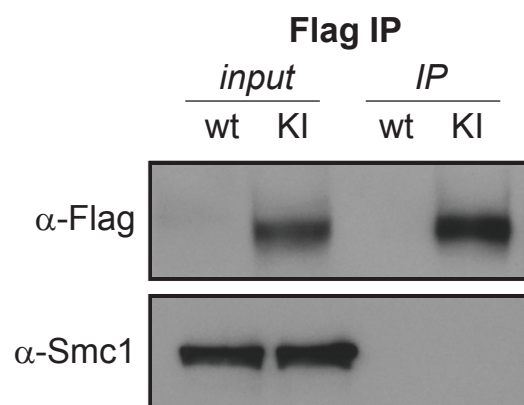
**Suv39h DKO
(D15)**

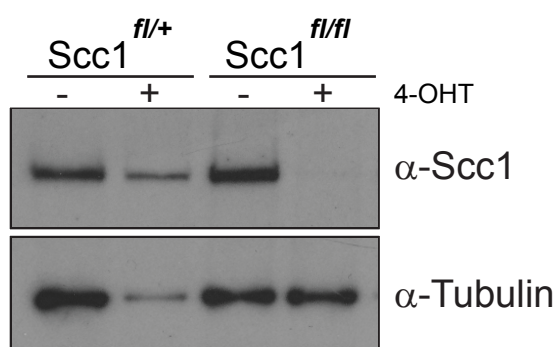
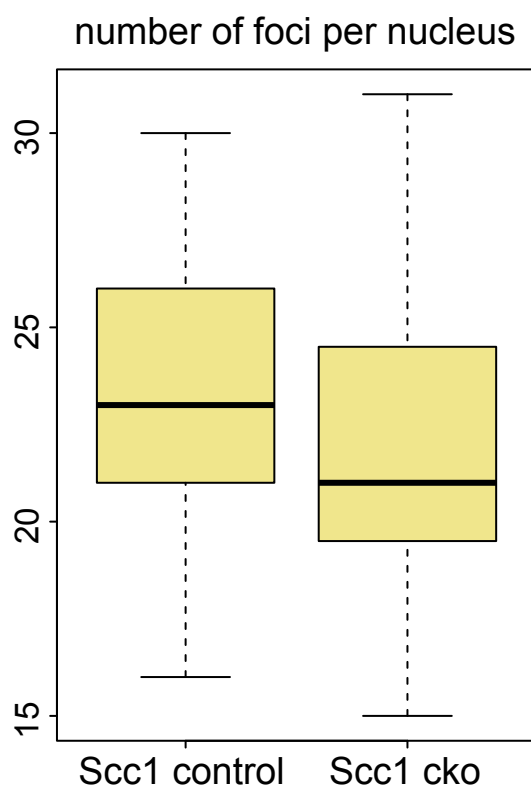








A**B****C**

A**B****Shapiro-Wilk test**

Scc1 control: W=0,9193
p-value=0,04929

Scc1 cKO: W = 0.8979
p-value = 0.01024

F-test

Scc1 control
vs. Scc1 cKO: p-value=0,04495

Welch-test

Scc1 control
vs. Scc1 cKO: p-value=0,1564

SUPPLEMENTAL DATA

Figure S1. FRAP analysis of heterochromatin core proteins.

(A) Heterochromatin is marked by two prominent histone modifications, H3K9me3 and H4K20me3. Suv39h enzymes are recruited by currently unknown mechanisms to pericentric heterochromatin and induce H3K9me3. This modification is then bound by HP1 proteins. Suv4-20h proteins bind to heterochromatin in an HP1-dependent manner and induce H4K20me3. (B) FRAP analysis in wild type fibroblast cells expressing HP1 α -EGFP or Suv4-20h2-EGFP. Red circle marks the bleached region which corresponds to pericentric heterochromatin.

Figure S2. Mobility of Suv4-20h2 at endogenous expression level

(A) *Suv4-20h2* knock-in strategy to express C-terminally tagged Suv4-20h2-EGFP from the endogenous locus. The region of the targeting construct is shown as bold line. After homologous recombination, exon 8 contains an insertion of the EGFP open reading frame just before the regular STOP codon, followed by a neomycin resistance cassette. (B) Western blot of *Suv4-20h2*^{EGFP} and wild type ES cells using an antibody specifically recognizing Suv4-20h2 (Dambacher et al. 2012). (C) Confocal microscopy section of *Suv4-20h2*^{EGFP} knock-in cells. The EGFP-tagged Suv4-20h2 is clearly enriched at pericentric heterochromatin, together with H4K20me3. (D) Quantification of the transiently bound pool Suv4-20h2 and HP1 isoforms via FCS in terms of the apparent diffusion coefficient. (E) Immobile fraction of Suv4-20h2 and HP1 β as measured by continuous photobleaching. (F) Western blot of wild type fibroblast cells which were transiently transfected with EGFP-tagged *Suv4-20h2* truncations. Due to variations in transfection efficiency the expression levels of the different truncations cannot be directly compared. The minus sign (-) denotes an empty lane.

Figure S3. *In vivo* HP1 interaction test.

BHK cells containing a *lac* operator repeat array were co-transfected with expression vectors for lacR-GBP, Suv4-20h2-EGFP and HP1 isoforms as mCherry fusions. The lacR-GBP fusion protein binds to the *lac* operator array and recruits Suv4-20h2-EGFP. All HP1 isoforms interact with Suv4-20h2 *in vivo* and show co-localization at the *lac* operator arrays. Histograms display the percentage of cells where HP1 proteins co-localize with EGFP and Suv4-20h2-EGFP, respectively.

Figure S4. HP1 α knock-out cells display normal heterochromatin marks.

(A) Wild type and HP1 α KO MEFs were stained for HP1 α , H3K9me3 and H4K20me3. HP1 α is not detectable in the HP1 α KO cells. H3K9me3 and H4K20me3 are enriched at heterochromatin in both wild type and mutant cells (scale bars, 5 μ m). (B) Localization of Suv4-20h2 truncation proteins in HP1 α KO cells. Confocal sections of HP1 α KO fibroblasts expressing EGFP-tagged Suv4-20h2 truncations are shown (scale bars, 5 μ m).

Figure S5. Suv4-20h2 mediates chromatin compaction.

(A) MNase accessibility assay using fibroblast cell lines. Nuclei from wild type, *Suv4-20h2* DKO and rescue cells expressing full length *Suv4-20h2* (DKO+FL) or *Suv4-20h2*-M12 (DKO+M12) were isolated and incubated with increasing amounts of MNase. The digested DNA was purified, separated on agarose gels and stained with ethidium bromide (upper panel). Heterochromatic regions were visualized by Southern blotting using a major satellite-specific probe (lower panel). (B) Overlay of DNA electropherograms obtained from Bioanalyzer runs of the MNase-digested DNA from wild type, *Suv4-20h2* DKO and rescue cells (left panel). The first and the last sharp peak represent the markers of the Agilent DNA 12000 kit (FU = fluorescence units). Intensity linescan of major satellite Southern blots (right panel). Gray values correspond to signal intensities of the southern blot.

Figure S6. Reduced peripheral heterochromatin and altered arrangement of nuclear pores in *Suv4-20h* DKO cells.

(A) 3D-SIM analysis of nuclear structure (DAPI) was performed in wild type, *Suv4-20h* DKO and wild type cells over-expressing *Suv4-20h2*. Inserts are enlargements of the nuclear periphery. (B) Nuclear periphery of central sections of the 3D-SIM images was automatically segmented to measure DAPI intensities. Example images for segmented wild type and *Suv4-20h* DKO nuclei are shown. The area highlighted in blue was used to measure the DAPI intensity of peripheral heterochromatin. (C) The relative intensities of peripheral heterochromatin were calculated over several 3D-SIM images as ratio of peripheral DAPI intensity vs. total nuclear DAPI intensity. The P values of Student's t-test are indicated. (D) Automated 3D image analysis of nuclear pores. 3D-SIM images were automatically segmented based on the DAPI channel to identify the nucleus and the anti-NUP staining to identify nuclear pores. Image shows an example wild type nucleus as 3D visualization of the segmentation result. Black spots – nuclear pores on the top surface; gray spots – nuclear pores on

the bottom surface. (E) Quantification of the nuclear pore segmentation. The average nuclear pore density is not significantly different between wild type and *Suv4-20h* DKO cells, or cells over-expressing *Suv4-20h2* (P values of Student's t-test). Regions with a low density of nuclear pores (sparse nuclear pore regions: less than 2 pores in a radius of 0.5µm) are less abundant in *Suv4-20h* DKO cells and are strongly increased in cells over-expressing *Suv4-20h2* (P values of Chi-square with Yates correction).

Figure S7. Large-scale chromatin compaction upon over-expression of Suv4-20h2.

Mouse embryonic fibroblast cells were transfected with constructs expressing EGFP or EGFP-tagged Suv4-20h2 truncations. 3D-SIM analysis of nuclear structure (DAPI) was then performed in EGFP-positive cells. Over-expression of *Suv4-20h2* full length and C-terminus (aa 280-468) lead to large-scale chromatin compaction (increased chromatin density of peripheral heterochromatin and around nucleoli). Neither the N-Terminus (aa1-280) of Suv4-20h2 nor the clamp domain (M12) lead to obvious alterations in nuclear structure.

Figure S8. 3D image segmentation of major satellite FISH experiments.

Example results of 3D nuclei segmentation (green) and 3D foci segmentation (red) for nuclei of three different cell types: wild type cells (first row), *Suv4-20h* DKO cells (second row), and *Suv4-20h* DKO cells which re-express *Suv4-20h2* (third row). The first and second column show contours of the 3D segmentation results overlaid with the original images (maximum intensity projections, MIPs) of the FISH and DAPI channel, respectively. The third column shows 3D visualizations of the segmentation results.

Figure S9. Statistical tests for 3D image segmentation results.

Results of statistical relevance tests performed on the quantified number *n* of foci per nucleus: (A) Shapiro-Wilk test testing the null hypothesis that *n* is normally distributed for wild type cells, *Suv4-20h* DKO cells and rescue cells (DKO+FL). For DKO+M12 the null hypothesis is rejected. Therefore a two-sample Wilcoxon Rank Sum Test was performed showing no significant differences between DKO and DKO+M12. (B) F-test for homogeneity of variances of *n* (null hypothesis) between different cell types, (C) Welch-test for homogeneity of means of *n* (null hypothesis) between

different cell types, indicating that the difference in FISH foci numbers is significantly different between wild type and *Suv4-20h* DKO cells.

Figure S10. The functional SET domain and the clamp domain of *Suv4-20h2* are important to mediate chromocenter clustering.

Quantification of FISH analyses of major satellite repeats in G0 phase *Suv4-20h* DKO and *Suv4-20h2* rescue cells expressing *Suv4-20h2* with deletion of the clamp domain (DKO+ Δ M12) or *Suv4-20h2* containing mutations that render the SET domain inactive (DKO+SET*). Chromocenter foci were counted and measured using an image segmentation analysis of confocal 3D stacks. Boxplots showing the numbers of chromocenters in *Suv4-20h* DKO (n=28), DKO+ Δ M12 (n=25) and DKO+SET* (n=25) nuclei. The mean number of foci per nucleus is not significantly different between the three cell lines (statistical tests which correspond to the analysis in Figure S9 are shown in the right panel).

Figure S11. Mitotic stages in wild type and *Suv4-20h* DKO MEFs.

Progression through mitosis is normal in *Suv4-20h* DKO cells. Different mitotic stages were counted for wild type (n=285) and *Suv4-20h* DKO (n=335) MEFs (P-prophase, PM/M-prometaphase/metaphase, A-anaphase, T-telophase).

Figure S12. *Suv4-20h2* GST pulldown and knock-in strategy.

(A) GST pulldown experiments were performed as described in materials and methods. Proteins which bound to the GST beads were eluted and loaded on a protein gel followed by silver staining. Recombinant GST, *Suv4-20h2*-M5 and *Suv4-20h2*-M7 proteins are marked with an arrow head, degradation products of the recombinant proteins are marked with asterisks. (B) PCR analysis for major satellite repeats with DNA isolated from sonicated extracts that were used for GST pulldown experiments and DNA isolated from benzonase-treated extracts. Undigested genomic DNA was used as control (gDNA). (C) Knock-in strategy for the *Suv4-20h2* locus. The region of the targeting construct which was used for homologous recombination in wild type ES cells is shown as thick line. The knock-in allele contains the coding sequence for a HA-3xFLAG tag immediately after the last codon of *Suv4-20h2*. A FRT-flanked neomycin selection cassette was used for positive selection of the knock-in clones. (D) The clamp domain in *Suv4-20h2* is important for cohesin interaction. Nuclear extract was prepared from wild type fibroblast cell lines expressing EGFP, *Suv4-20h2*-EGFP or *Suv4-20h2*-DM12-EGFP. *Suv4-20h2* was

precipitated from these extracts using EGFP affinity beads. Bound proteins were visualized by western blotting using Suv4-20h2 and Smc1 antibodies.

Figure S13. Cohesin ChIP in wild type, Suv4-20h2 KO and Suv4-20h DKO cells.

ChIP analysis for histone modifications (H3K9me3, H4K20me3) and cohesin subunits (Rad21, Smc3) in wild type (E19-8, W9), *Suv4-20h2* KO (E29-7, E27-5) and *Suv4-20h* DKO (E94-4, E35-3) cell lines. Major satellite repeats represent pericentric heterochromatin. Control regions where cohesin is enriched independently from H3K9me3-H4K20me3 (pos controls: Chr.6, Chr.7, Chr.10) and where cohesin is not present (neg. controls: Chr.5, Chr. 6-1, Chr.10-1) are shown. Merge of two experiments with three technical replicates, each. Error bars represent standard deviation.

Figure S14. Cohesin ChIP in wild type and Suv39h DKO cells.

ChIP analysis for histone modifications (H3K9me3, H4K20me3) and the cohesin subunit Smc3 in wild type (W9) and two independent *Suv39h* DKO cell lines. Major satellite repeats represent pericentric heterochromatin. Control regions where cohesin is enriched independently from H3K9me3-H4K20me3 (pos controls: Chr.6, Chr.7, Chr.10) and where cohesin is not present (neg. controls: Chr.5, Chr. 6-1, Chr.10-1) are shown. Merge of three technical replicates (two independent experiments for E124-1). Error bars represent standard deviation.

Figure S15. Cohesin ChIP in wild type and Suv4-20h rescue cells.

ChIP analysis for histone modifications (H3K9me3, H4K20me3) and cohesin (Smc3) were performed in wild type cells and *Suv4-20h* DKO cells which express full length Suv4-20h2-EGFP (DKO+FL) or Suv4-20h2-M12 (DKO+M12). The control region Chr.6 displays detectable levels of H4K20me3 upon re-expression of Suv4-20h2 in DKO cells, which is probably an over-expression effect. Merge of two experiments with three technical replicates, each. Error bars represent standard deviation.

Figure S16. FACS analyses of cell cycle synchronized fibroblasts.

FACS analysis of cells in G0, S and M phase. Cells were isolated as described and stained with propidium iodide prior to FACS analysis.

Figure S17. Cohesin ChIP in wild type and Suv4-20h rescue cells in different cell cycle stages.

ChIP analysis for histone modifications (H3K9me3, H4K20me3) and the cohesin subunit Smc3 were performed in wild type, *Suv4-20h* DKO and rescue cell lines expressing full length Suv4-20h2 (DKO+FL) or the Suv4-20h2 clamp domain (DKO+M12).

Figure S18. Suv4-20h1 does not interact with cohesin.

(A) Suv4-20h1 is a heterochromatin-associated protein. EGFP-tagged full length (C0) and different truncations of Suv4-20h1 were expressed in wild type fibroblast cells. Fragment C4 is the smallest truncation which is still able to localize to heterochromatin. This truncation contains one classical PxVxL HP1 interaction motif. (B) Knock-in strategy for the *Suv4-20h1* locus. The region of the targeting construct which was used for homologous recombination in wild type ES cells is shown as thick line. The knock-in allele contains the coding sequence for a HA-3xFLAG tag immediately after the last codon of *Suv4-20h1*. A FRT-flanked neomycin selection cassette was used for positive selection of the knock-in clones. (C) Suv4-20h1 does not interact with cohesin. Nuclear extracts were prepared from wild type (wt) and *Suv4-20h1*^{HA-FLAG} knock-in ES cells (KI). Suv4-20h1 was precipitated from these extracts using a FLAG antibody. Bound proteins were visualized by western blotting using FLAG and Smc1 antibodies.

Figure S19. Cohesin is not important to mediate chromocenter clustering.

(A) Inducible deletion of *Sccl*/Rad21. *Sccl*^{fl/+}; Cre-ERT2 and *Sccl*^{fl/fl}; Cre-ERT2 (cko) cells (Seitan et al. 2011). Fibroblasts were arrested in G0 by serum deprivation. Addition of 4-OHT activates the Cre recombinase which leads to conversion of the *Sccl* flox to a *Sccl* mutant allele. Three weeks after addition of 4-OHT total cell extracts were prepared and analyzed for *Sccl* protein levels compared to tubulin loading control. In 4-OHT-induced *Sccl* cko cells *Sccl* protein is not detectable. The reduced signal intensity in lane 2 (induced *Sccl*^{fl/+} cells) is due to reduced loading. (B) Quantification of FISH analyses of major satellite repeats in G0 phase non-induced (control) vs. 4-OHT induced *Sccl* cko cells. Chromocenter foci were counted and measured using an image segmentation analysis of confocal 3D stacks. Boxplots showing the numbers of chromocenters in *Sccl* control (n=25) and *Sccl* cko (n=28) nuclei. The mean number of foci per nucleus is not significantly different

between the two cell lines (statistical tests which correspond to the analysis in Figure S9 are shown in the right panel).

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Cloning of Suv4-20h2-ΔM12:

Internal deletion mutants of the *Suv4-20h2* cDNA were created by inverse PCR on the *Suv4-20h2* wild type plasmid pDONR207-Suv4-20h2 using a reverse primer upstream and a forward primer downstream of the M12 region. Primer overhangs were designed to create 15 bp complementary ends on both sides of the PCR product to allow seamless recircularization using the Infusion HD cloning kit (Clontech).

Cell culture

MEF and ES cells were cultivated in High Glucose DMEM with L-Glutamine complemented with sodium pyruvate, fetal calf serum, beta-mercaptoethanol, non-essential amino acids (PAA) and penicillin/streptomycin (PAA) in a 37°C incubator at 5% CO₂. For ES cell culture medium was supplemented with leukemia inhibitory factor (LIF). Cells were transfected using Lipofectamine 2000 (Invitrogen) or TransIT-LT1 (Mirus). ES cells were seeded on matrigel (BD Biosciences) coated slides 24h before FRAP analysis.

FRAP

FRAP measurements were conducted in MEF and ES cell lines after transient expression of EGFP-tagged proteins using a Leica TCS SP5 confocal laser scanning microscope (CLSM) equipped with a HCX PL APO lambda blue 63x/1.4 NA oil immersion objective or an UltraVIEW VoX spinning disk microscope system (PerkinElmer). The region of interest (ROI) with an effective diameter of 1.9 μm was subjected to high intensity laser pulses to irreversibly bleach EGFP. Post-bleach images were collected at different time intervals with the laser intensity attenuated to the same level as in the pre-bleach images. Analysis of the recovery curves of the intensity integrated over the ROI was done using FREDIS software (Muller et al. 2009). For Figure 1 the averaged recovery curves have been fitted to a diffusion model or a reaction-diffusion model that incorporates both diffusion and binding processes.

Fluorescence Correlation Spectroscopy

For FCS experiments a Zeiss LSM 710 ConfoCor3 was used. Before measurements, the system was calibrated with Alexa Fluor 488 C maleimide (Invitrogen) and from

these measurements the effective focal volume was determined. Similar laser intensities were used for all experiments with an acquisition time of 60 s. For each FCS measurement a continuous photobleaching (CP) fit was applied to determine the immobile fraction as described previously (Wachsmuth et al. 2003; Muller et al. 2009). FCS data were analyzed as previously described (Erdel et al. 2010) with minor alterations. Since bleaching occurred during the measurements, especially in heterochromatin regions, only intensity fluctuations that were monitored 10 s after starting the acquisition were used for fittings. Then a moving average trend correction was applied and an autocorrelation function was fitted for anomalous diffusion. Over 15 individual cells per genotype were measured and averaged to derive the mobility data from the FCS based experiments. FCS data for HP1 were taken from (Muller et al. 2009).

HP1 interaction test

GST fusion proteins of Suv4-20h2, HP1 α , HP1 β and HP1 γ were expressed in *E. coli* and purified using Glutathione-S-Sepharose beads (Amersham Biosciences). GST-tags of HP1 α , HP1 β and HP1 γ were removed by PreScission (GE Healthcare) cleavage. Interaction tests were performed by incubating 5 μ g GST-Suv4-20h2 fusion protein with 5 μ g of either HP1 α , HP1 β and HP1 γ in IP buffer [50 mM Tris pH 7.5, 150 mM NaCl, 1 mM EDTA, 0.1% NP40, 20% glycerol and proteinase inhibitor cocktail (Roche)] for 1h at room temperature on a rotating wheel. Then 60 μ l Glutathione-S-Sepharose beads were added and incubated for 1h at 4°C on a rotating wheel. The beads were washed 3x with IP buffer containing 300 mM NaCl and eluted with 50 μ l SDS sample buffer. Proteins were separated on SDS-polyacrylamide gels and detected by immunoblotting using isoform-specific HP1 antibodies (Euromedex).

Chromatin accessibility assay

Cytoplasm of wild type, Suv4-20h2 KO and Suv4-20h2 DKO ES cells and wild type, Suv4-20h DKO, DKO+FL and DKO+M12 fibroblast cell lines was removed by spinning through a Ficoll gradient. Isolated nuclei were resuspended in Ex100 buffer [10 mM HEPES pH 7.6, 100 mM KCl, 1.5 mM MgCl₂ 0.5 mM EGTA, 10% Glycerol, proteinase inhibitor cocktail (Roche)] and digested with different amounts of MNase (Sigma) in the presence of 3mM CaCl₂. Digest was stopped after 9 min by addition of 1/10 volume MNase stop buffer [0.5 M EDTA, 10% SDS]. Genomic DNA was purified and then separated on a 1% agarose gel, followed by blotting onto a Nylon membrane [Roti-Nylon plus, Roth] via capillary transfer. Hybridization with a ³²P-

labeled major satellite probe was performed using standard conditions. MNase digestion patterns were further analyzed using the Agilent 2100 bioanalyzer (Agilent Technologies). 100ng of each sample were loaded onto a DNA LabChip using the Agilent DNA 12000 kit and analysis was performed using the 2100 expert software (Agilent Technologies). Southern blot quantification was performed using the line selection and plot profile function of ImageJ.

Lac-recruitment assay

EGFP-tagged proteins were recruited via a lac I repressor-GFP binding domain fusion protein to telomeres and promyelocytic leukemia nuclear bodies in a telomerase-negative U2OS cell carrying three integrations of lacO arrays as described (Jegou et al. 2009). 24 h after transfection, cells were fixed and mounted with Prolong (Invitrogen). Images were taken with a Leica SP5 microscope and pictures were analyzed with ImageJ. Signal intensity of mCherry was measured at lacO arrays and regions adjacent to the lacO arrays. Positive recruitment was defined as the signal increase at the lacO array by >1.5 fold as compared to adjacent regions.

3D-structured illumination microscopy

Fibroblasts were seeded on No. 1.5H precision coverslips (thickness 0.170 ± 0.005 mm; Marienfeld Superior), fixed with 4% formaldehyde, permeabilized with 0.5% Triton X-100, stained with 2 $\mu\text{g/ml}$ 4',6-diamidino-2-phenylindole (DAPI) and embedded in Vectashield mounting medium (Vector Labs). Nuclear pore complexes (NPCs) were immunostained with monoclonal anti-Nup153 antibodies (QE5, Abcam) and Alexa Fluor 594 conjugated secondary antibodies (Invitrogen). Super-resolution 3D-structured illumination microscopy (3D-SIM) was performed as previously described (Schermele et al. 2008) on a Delta Vision OMX V3 system (Applied Precision) equipped with a 100x /1.40 NA PlanApo oil immersion objective (Olympus) and Cascade II:512 EMCCD cameras (Photometrics), using 405 nm and 592 nm diode lasers. In brief, a fine-striped interference pattern of illumination was generated by coherent laser light directed through a movable optical grating. Raw image stacks with 15 images per plane (5 phase shifts, 3 angles) and a z-distance of 125 nm were acquired and computationally reconstructed (SoftWoRX, Applied Precision) to obtain a super-resolution image dataset with a two-fold resolution improvement in x, y and z direction (eight-fold volumetric resolution improvement) compared to conventional diffraction-limited optical sectioning microscopy (Gustafsson et al. 2008). Images

from the different color channels were registered with alignment parameter obtained from calibration measurements with 0.2 μm diameter TetraSpeck beads (Invitrogen).

Mitotic abnormalities

To quantify the number of abnormal mitotic figures in MEF cells, mitotic cells were collected by shake-off, spun onto glass-bottomed 96-black well plate (Greiner Bio-One) for 5min at 400g, fixed in 4% formaldehyde and stained with Vectashield containing DAPI.

To quantify mitotic stages MEF cells were grown for 48h cells on cover slips and fixed with 4% formaldehyde. After mounting with Vectashield containing DAPI cells were classified according to DAPI stain and mitotic spindle morphology into different stages: P-prophase; PM/M–prometaphase/metaphase; A-anaphase; T-telophase.

Immunofluorescence analysis

Immunofluorescence analyses were performed as described (Dambacher et al. 2012) using the following antibodies: CENP-A (C51A7, Cell Signaling Technology), AuroraB (Abcam ab2254), H4K20me3 (Schotta et al. 2008), Sgo1 (Whelan et al. 2012), Nup-153 (QE5, Abcam) and secondary antibodies conjugated to Cy5 (Dianova).

For chromosome spreads, untreated or nocodazole arrested (final concentration 200 nM, 4 h) mitotic cells were collected by shake-off, swollen for 10 min at room temperature in 75 mM KCl and spun onto microscope slides using a Cytospin centrifuge (Fischer). Cells were extracted for 5 min at room temperature with KCM-buffer [120 mM KCl, 20 mM MgCl_2 , 10 mM, Tris pH 8.0, 0.5 mM EDTA, 0.1% Triton X-100] and fixed in 2% formaldehyde. Slides were blocked for 30 min in KCM/0.1% Tween-20/2.5% BSA, and sequentially incubated with primary and secondary antibodies. After washing with KCM/0.1% Tween-20 chromosomes were mounted with Vectashield containing DAPI.

A Leica TCS SP5 confocal laser scanning microscope with a HCX PL APO CS 63x/1.3 NA glycerol immersion objective and a Zeiss LSM510 confocal microscope were used to obtain the images. Sequential excitation at 405 nm, 488 nm, 543 nm and 633 nm was provided by diode, argon and helium-neon gas lasers, respectively. Emission detection ranges of the PMTs were adjusted to minimize crosstalk between the channels. Images were further analyzed using ImageJ software.

HP1 knock-out cell line

The W205C07 gene trap ESC line carrying a null mutation in the first intron of the HP1 α gene was isolated from a large scale genetrap screen (Hansen et al. 2003) carried out in 129S2/SvPAS ES cells with the retroviral gene trap vector Rosabgeo. The W205C07 ESCs were converted into mice by standard procedures using C57BL6J host blastocysts. The 129/C57BL6J F1 offspring were back-crossed on C57BL6J for over 10 generations. The HP1 α homozygous mutants are viable and show no obvious phenotype (PD and HvM, unpublished). Mouse Embryonic Fibroblasts (MEFs) were derived by standard procedures from E13.5 embryos and immortalized using the 3T3 protocol.

GST pull-down assays

GST fusions of Suv4-20h2 truncations were expressed from pGEX6P1 vector in *E. coli* and purified using Glutathione-S-Sepharose beads (Amersham Biosciences). Nuclei of wild type ES cells were isolated by spinning through a Ficoll gradient. For the Benzonase treatment cells were resuspended in low salt IP buffer [50 mM Tris pH 7.5, 150 mM KCl, 1 mM EDTA, 20% glycerol, 0.1% NP40] and digested with Benzonase (Merck Chemicals) for 15 min at 30°C, then adjusted to 500 mM KCl, incubated on ice for 30 min followed by 3x 10 sec sonication at an amplitude of 30 in a Branson sonifier. For the salt extract, proteins were extracted with high salt IP buffer [50 mM Tris pH 7.5, 500 mM KCl, 1 mM EDTA, 20% glycerol, 0.1% NP40, proteinase inhibitor cocktail (Roche)] incubated on ice for 30 min, followed by 3x 10 sec sonication at an amplitude of 30 in a Branson sonifier. Both nuclear extracts were diluted to a final concentration of 250 mM KCl with no salt IP buffer; precipitated or non-soluble proteins were then removed by centrifugation. 5 μ g of either GST or the different GST-Suv4-20h2 fusion proteins were incubated with the Benzonase treated nuclear extract or salt extract supplemented with 50 μ g/ml EtBr over night at 4°C on a rotation wheel. Precipitates were again removed by centrifugation before 60 μ l Glutathione-S-Sepharose beads were added and incubated for another hour at 4°C. Beads were washed with IP buffer containing 300 mM KCl; subsequently, the bound proteins were eluted with 50 μ l SDS sample buffer. Proteins were separated on a SDS-polyacrylamid gel, visualized by silver staining and identified by mass-spectrometry analysis.

Co-immunoprecipitation

Nuclei of wild type and *Suv4-20h2*^{HA-FLAG} ES cells were isolated by spinning through a Ficoll gradient. The nuclear pellet was resuspended in low salt IP Buffer [50mM Tris pH 7.5, 150mM KCl, 1mM EDTA, 20% glycerol, 0.1% NP40, proteinase inhibitor cocktail (Roche)]. Fraction I extract was isolated after digestion for 15 min at 37°C with Benzonase (Merck) in low salt IP buffer, followed by mild sonication 1x 10 sec at amplitude 20 in a Branson sonifier. Insoluble proteins were separated by spinning for 10 min at 13000 rpm at 4°C. The protein pellet was further extracted with high salt IP Buffer [50mM Tris pH 7.5, 500mM KCl, 1mM EDTA, 20% glycerol, 0.1% NP40, proteinase inhibitor cocktail (Roche)] on ice for 30 min followed by 3x 10 sec sonication at an amplitude of 30 in a Branson sonifier. Fraction II extract was then diluted to a final concentration of 250mM NaCl with no salt IP Buffer and precipitated or insoluble proteins were removed by centrifugation (30min, 13000rpm, 4°C). The different extracts were incubated with either 5µg Flag M2 antibody (Sigma) or Smc1 antibody (Bethyl Labs) coupled to ProteinA/G magnetic beads (Dynabeads Invitrogen) o/n at 4°C on a rotating wheel. The beads were washed 5x in IP buffer containing 300mM KCl and eluted with SDS sample buffer. Proteins were separated on SDS-polyacrylamide gels and analyzed by western blotting using antibodies for Flag M2 (Sigma), Smc1 (Bethyl labs), Smc3 (Abcam), Suz12 (Cell Signalling) and Suv4-20h2 (Dambacher et al. 2012). Immunoprecipitation experiments of EGFP-tagged proteins were performed using GFP-Trap A beads (ChromoTek).

Chromatin immunoprecipitation (ChIP)

ChIP experiments were performed according to standard methods. Chromatin was sonicated to an average size of 300-600 bp. 50 µl Dynabeads Protein A/G beads (Dynabeads, Invitrogen) were pre-bound with 3-5 µg antibody (anti-H3K9me3 Abcam #ab8898, anti-H4K20me3 Active Motif #39180.3918 or Millipore #07-463, anti-Rad21 Abcam #ab992, anti-Smc3 Abcam #ab9263) and added to the chromatin. After 16 h incubation at 4°C on a rotating wheel beads were collected by DynaMag magnet (or Dynal MPC) for 1 min. Beads were washed with Sonication buffer [50 mM Hepes pH 7.9, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-deoxycholate, 0.1% SDS, 0.5 mM PMSF, Protease inhibitor cocktail (Roche)], wash buffer A [Sonication buffer supplemented with 500mM NaCl], wash buffer B [20 mM Tris, pH 8.0, 1 mM EDTA, 250 mM LiCl, 0.5% NP-40, 0.5% Na-deoxycholate, 0.5 mM PMSF, Roche protease inhibitor cocktail] and TE buffer pH 8.0 [10 mM Tris-Cl pH 8.0, 1 mM EDTA pH 8.0 (supplemented with 50 mM NaCl)]. 200 µl of elution buffer [50 mM Tris pH

8.0, 1 mM EDTA, 1% SDS, 50mM NaHCO₃] was added to the beads for 20 min at 65°C in two rounds. Eluates were purified after crosslink reversal using the GenElute PCR Clean-Up Kit (Sigma).

Cell cycle arrest

MEF cells were plated onto 150 mm plates at 40-50% confluency and treated with 2 mM thymidine. After 18 h cells were washed 3x with medium, and after 8 h the second thymidine block was applied. Cells were released after 17 h by 3 washes with medium and harvested 3 h after the release. For mitotic arrest MEF cells were first treated with thymidine (2 mM) for 20h. Then cells were washed 3x with medium, and after 3 h nocodazole was added to 100 ng/ml. Cells were harvested 11 h after the thymidine release by mitotic shake-off. Cells were arrested in G0 by 48 h serum starvation. The *Sccl/Rad21* conditional knock-out cell line (*Sccl*^{lox/lox}; Cre-ERT2) was arrested in G0 by serum starvation and after 72 h treated with ethanol (*Sccl* control) or induced with 250 nM final concentration of 4-OHT (solubilized in ethanol) for 21 days (*Sccl* cko). The medium was exchanged every 3 days.

FACS

Cells were stained with propidium iodide (PI) to analyze the cell cycle phase by DNA amount. Trypsinized MEF cells were fixed with 70% methanol for 30 min on ice and PBS washed. PI staining was performed at 37°C for 30 min in a buffer containing 10 mM Tris pH 7.5, 5 mM MgCl₂ and 200 µg RNase A.

Suv4-20h1 and Suv4-20h2 knock-in cell lines

The *Suv4-20h1*^{HA-FLAG} and *Suv4-20h2*^{HA-FLAG} targeting constructs were obtained using the recombineering cloning technique described previously. The NotI-linearized targeting vector was electroporated into feeder-independent wild type ES cells. Cells were selected in 180 µg/ml G418 (PAA) and 2 µM Ganciclovir (Invivogen). Single colonies were picked and screened by nested PCR to obtain the final ES cell clones. The same strategy was used for generating *Suv4-20h2*^{EGFP} knock-in ES cells, except that the targeting construct contained the EGFP open reading frame fused to the last codon of *Suv4-20h2*. Primers used for cloning and PCR screening are listed in Table S1.

Automatic 3D Quantification of FISH Foci

To determine quantitative information of FISH major satellite repeats from the acquired two-channel 3D confocal microscopy images, we developed a fully automatic 3D image analysis approach consisting of 3D cell nuclei segmentation and 3D foci segmentation. First, 3D cell nuclei segmentation is performed by Otsu thresholding based on the information in the DAPI channel. This enables to determine relevant parameters (e.g., number and mean volume of FISH foci) for each nucleus individually. To ensure a valid statistical analysis, cell nuclei at the image border which are only partially visible are automatically excluded using a threshold on the number of border voxels of the segmented nucleus volume. Second, after cell nuclei segmentation, 3D foci segmentation is performed in the FISH channel using global thresholding and morphological opening. As a result, quantitative information of FISH foci (e.g., volume, 3D position, and 3D shape) is obtained. In addition, the 3D regions of segmented FISH foci are used to quantify heterochromatin concentrations in the DAPI channel.

In a subsequent statistical analysis, we further investigated the mean number of foci per nucleus denoted by n and the quantified foci volumes denoted by v . We subdivided the results into 3 groups w.r.t. the experimental setup: (1) wild type cells, (2) *Suv4-20h* DKO cells, (3) re-express *Suv4-20h2* cells. To ensure that our findings (Fig. 5c) are statistically relevant, we first performed a Shapiro-Wilk test, which indicated that in all groups n is normally distributed, and an F-Test which revealed that the variance of n differs significantly for groups (1) vs. (3). Therefore, we used a Welch-test which showed that the mean values of n for groups (1) vs. (2) and (3) vs. (2) differ significantly, while the mean values of n of groups (1) vs. (3) do not differ significantly. To investigate the foci volumes, we computed histograms of v for each group. It can be seen, that group (2) contains significantly more small foci with a volume < 2500 voxels, while groups (1) and (3) contain more foci with a larger volume.

Quantification of peripheral heterochromatin and nuclear pores

To determine quantitative information about the nuclear pores from the two-channel 3D SIM images, we developed an automatic 3D image analysis approach consisting of 3D cell nuclei segmentation and 3D model-based segmentation of the pores. First, the 3D cell nuclei are segmented in the DAPI channel by multi-level 3D Otsu thresholding and by using 3D grayscale morphological filters. Second, the 3D nuclear pores are localized and segmented in the NPC channel using an automatic 3D

approach based on a 3D anisotropic Gaussian parametric intensity model (Worz et al. 2010), which is accurate for small subcellular structures. As a result, quantitative information about each of the segmented nuclear pores (e.g., volume, 3D position, and distance to neighboring pores) is obtained with subvoxel accuracy. By using the results of 3D cell nuclei segmentation, relevant parameters like, e.g., the mean volume or mean distance to neighboring pores, are determined individually for each cell nucleus.

To quantify the amount of heterochromatin near the nuclear lamina, 2D slices along the z-axis from the segmented 3D cell nuclei volumes are extracted. For each extracted slice, a region-of-interest is defined near the nuclear lamina within a 150 nm radius from the contour of the segmented nucleus. Then, within the region-of-interest the total amount of signal intensity in the DAPI channel is determined.

SUPPLEMENTAL REFERENCES

- Dambacher S, Deng W, Hahn M, Sadic D, Frohlich J, Nuber A, Hoischen C, Diekmann S, Leonhardt H, Schotta G. 2012. CENP-C facilitates the recruitment of M18BP1 to centromeric chromatin. *Nucleus* **3**.
- Erdel F, Schubert T, Marth C, Langst G, Rippe K. 2010. Human ISWI chromatin-remodeling complexes sample nucleosomes via transient binding reactions and become immobilized at active sites. *Proc Natl Acad Sci U S A* **107**: 19873-19878.
- Gustafsson MG, Shao L, Carlton PM, Wang CJ, Golubovskaya IN, Cande WZ, Agard DA, Sedat JW. 2008. Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination. *Biophysical journal* **94**: 4957-4970.
- Hansen J, Floss T, Van Sloun P, Fuchtbauer EM, Vauti F, Arnold HH, Schnutgen F, Wurst W, von Melchner H, Ruiz P. 2003. A large-scale, gene-driven mutagenesis approach for the functional analysis of the mouse genome. *Proc Natl Acad Sci U S A* **100**: 9918-9922.
- Jegou T, Chung I, Heuvelman G, Wachsmuth M, Gorisch SM, Greulich-Bode KM, Boukamp P, Lichter P, Rippe K. 2009. Dynamics of telomeres and promyelocytic leukemia nuclear bodies in a telomerase-negative human cell line. *Mol Biol Cell* **20**: 2070-2082.
- Muller KP, Erdel F, Caudron-Herger M, Marth C, Fodor BD, Richter M, Scaranaro M, Beaudouin J, Wachsmuth M, Rippe K. 2009. Multiscale analysis of dynamics

- and interactions of heterochromatin protein 1 by fluorescence fluctuation microscopy. *Biophysical journal* **97**: 2876-2885.
- Schermelleh L, Carlton PM, Haase S, Shao L, Winoto L, Kner P, Burke B, Cardoso MC, Agard DA, Gustafsson MG et al. 2008. Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy. *Science* **320**: 1332-1336.
- Schotta G, Sengupta R, Kubicek S, Malin S, Kauer M, Callen E, Celeste A, Pagani M, Opravil S, De La Rosa-Velazquez IA et al. 2008. A chromatin-wide transition to H4K20 monomethylation impairs genome integrity and programmed DNA rearrangements in the mouse. *Genes Dev* **22**: 2048-2061.
- Seitan VC, Hao B, Tachibana-Konwalski K, Lavagnolli T, Mira-Bontenbal H, Brown KE, Teng G, Carroll T, Terry A, Horan K et al. 2011. A role for cohesin in T-cell-receptor rearrangement and thymocyte differentiation. *Nature* **476**: 467-471.
- Wachsmuth M, Weidemann T, Muller G, Hoffmann-Rohrer UW, Knoch TA, Waldeck W, Langowski J. 2003. Analyzing intracellular binding and diffusion with continuous fluorescence photobleaching. *Biophysical journal* **84**: 3353-3363.
- Whelan G, Kreidl E, Wutz G, Egner A, Peters JM, Eichele G. 2012. Cohesin acetyltransferase Esco2 is a cell viability factor and is required for cohesion in pericentric heterochromatin. *EMBO J* **31**: 71-82.
- Worz S, Sander P, Pfannmoller M, Rieker RJ, Joos S, Mechttersheimer G, Boukamp P, Lichter P, Rohr K. 2010. 3D geometry-based quantification of colocalizations in multichannel 3D microscopy images of human soft tissue tumors. *IEEE transactions on medical imaging* **29**: 1474-1484.

Table S1. Mass-spec list of the Suv4-20h2 GST pulldown assays. The number of unique spectra that were identified for each protein in the mass-spec analysis is indicated (Scaffold analysis).

gene	name	GS	M5	M7	delt
Suv420	histone-lysine N-methyltransferase SUV420H2	4	112	62	108
Lmnb1	lamin-B1	1	82	75	81
Top2a	DNA topoisomerase 2-alpha	0	49	0	49
Cbx5	chromobox protein homolog 5	0	41	0	41
Hnrnpu	heterogeneous nuclear ribonucleoprotein U	2	43	4	41
CBX1	chromobox protein homolog 1	0	40	0	40
Plec	plectin	0	40	8	40
CBX3	chromobox protein homolog 3	1	40	0	39
Jup	junction plakoglobin	4	37	13	33
Mybbp1	myb-binding protein 1A	0	30	0	30
Ahnak	AHNAK nucleoprotein	4	34	13	30
Krt31	keratin, type I cuticular Ha1	0	29	0	29
Ddx21	nucleolar RNA helicase 2	0	28	0	28
Nop58	nucleolar protein 58	0	23	0	23
Ncl	nucleolin	1	24	1	23
Nop56	nucleolar protein 56	0	23	12	23
543042	type II hair keratin	1	20	2	19
Eef2	elongation factor 2	0	19	8	19
Hnrnpa	heterogeneous nuclear ribonucleoprotein A3	4	23	12	19
Ddx5	probable ATP-dependent RNA helicase DDX5	4	22	1	18
Cltc	clathrin heavy chain 1	0	18	3	18
Vcl	vinculin	0	18	4	18
Acly	ATP-citrate synthase	0	18	5	18
Nup93	nuclear pore complex protein Nup93	0	16	0	16
Fasn	fatty acid synthase	0	16	3	16
Nolc1	nucleolar and coiled-body phosphoprotein 1	0	15	0	15
Hsp90b	endoplasmin	2	17	8	15
Nup107	nuclear pore complex protein Nup107	0	14	0	14
Tpr	nuclear pore complex-associated protein Tpr	0	14	0	14
Nedd4	E3 ubiquitin-protein ligase NEDD4	0	14	3	14
Ddx3x	ATP-dependent RNA helicase DDX3X	2	15	0	13
Smarca	SWI/SNF-related matrix-associated actin-dependent	0	12	0	12
Ranbp2	E3 SUMO-protein ligase RanBP2	0	12	3	12
Vcp	transitional endoplasmic reticulum ATPase	2	14	6	12
Spna2	spectrin alpha chain, brain	0	12	6	12
Vdac2	voltage-dependent anion-selective channel protein 2	3	14	1	11
Myh9	myosin-9	0	11	3	11
Atrx	transcriptional regulator ATRX	0	10	0	10
Ddost	dolichyl-diphosphooligosaccharide--protein	0	10	0	10
Las1l	protein LAS1 homolog	0	10	0	10
Nup85	nuclear pore complex protein Nup85	0	10	0	10
Nup98	nucleoporin 98	0	10	0	10
Snrnp20	U5 small nuclear ribonucleoprotein 200 kDa helicase	0	10	0	10
Ephx1	epoxide hydrolase 1	4	14	1	10
Vars	valyl-tRNA synthetase	0	10	2	10
Col1a1	collagen alpha-1(I) chain	2	12	4	10
Actn1	alpha-actinin-1	2	12	6	10
Hnrnpul	heterogeneous nuclear ribonucleoprotein U-like protein	0	9	1	9
Trim28	transcription intermediary factor 1-beta	0	9	1	9
Eef1b2	elongation factor 1-beta	3	12	3	9
Spnb2	spectrin beta chain, brain 1	0	9	4	9
Krt77	keratin, type II cytoskeletal 1b	4	13	6	9
Nat10	N-acetyltransferase 10	0	8	0	8
Orc1	origin recognition complex subunit 1	0	8	0	8

Pwp2	periodic tryptophan protein 2 homolog	0	8	0	8
Eprs	bifunctional aminoacyl-tRNA synthetase	0	8	1	8
Nup205	nucleoporin 205	0	8	1	8
Canx	calnexin	2	10	5	8
Krt78	keratin Kb40	4	12	6	8
Hist1h3	histone H3.2	1	9	9	8
Rif1	telomere-associated protein RIF1	1	8	0	7
Mcm3	DNA replication licensing factor MCM3	0	7	0	7
Tcof1	treacle protein	0	7	0	7
Vdac1	voltage-dependent anion-selective channel protein 1	3	10	1	7
Eif3c	eukaryotic translation initiation factor 3 subunit C	0	7	1	7
Smc1a	structural maintenance of chromosomes protein 1A	0	7	1	7
Rpl5	60S ribosomal protein L5	2	8	0	6
Mki67	antigen KI-67	0	6	0	6
Ptbp1	polypyrimidine tract-binding protein 1	0	6	0	6
Ruvbl1	ruvB-like 1	0	6	0	6
Sf3b1	splicing factor 3B subunit 1	0	6	0	6
Tmpo	lamina-associated polypeptide 2	0	6	0	6
Uqcrc1	cytochrome b-c1 complex subunit 1, mitochondrial	0	6	0	6
Alb	serum albumin	2	8	2	6
Ndufs1	NADH-ubiquinone oxidoreductase 75 kDa subunit,	0	6	2	6
Top2b	DNA topoisomerase 2-beta	2	8	3	6
Rrbp1	ribosome-binding protein 1	0	6	3	6
Snd1	staphylococcal nuclease domain-containing protein 1	0	6	3	6
Stt3a	dolichyl-diphosphooligosaccharide--protein	0	6	3	6
Aldh18a	delta-1-pyrroline-5-carboxylate synthase	0	6	4	6
LOC100		0	6	4	6
Krt75	keratin, type II cytoskeletal 75	1	7	5	6
Insrr	insulin receptor-related protein	2	8	7	6
Krt13	keratin, type I cytoskeletal 13	0	6	7	6
Hnrnpl	heterogeneous nuclear ribonucleoprotein L	2	7	0	5
Ddx17	probable ATP-dependent RNA helicase DDX17	0	5	0	5
Mcm5	DNA replication licensing factor MCM5	0	5	0	5
Numa1	nuclear mitotic apparatus protein 1	0	5	0	5
Otc	ornithine carbamoyltransferase, mitochondrial	0	5	0	5
Pdcd11	protein RRP5 homolog	0	5	0	5
Sf3a1	splicing factor 3A subunit 1	0	5	0	5
Sf3b2	splicing factor 3b, subunit 2	0	5	0	5
Wdr3	WD repeat-containing protein 3	0	5	0	5
Ddx3y	ATP-dependent RNA helicase DDX3Y	2	7	1	5
Lyz1	lysozyme C-1	3	8	2	5
Sfn	14-3-3 protein sigma	0	5	2	5
Dnmt3b	DNA (cytosine-5)-methyltransferase 3B	0	5	3	5
Hnrnpa	heterogeneous nuclear ribonucleoproteins A2/B1	4	9	6	5
Fam161	protein FAM161B	1	5	0	4
261010	U2 snRNP-associated SURP motif-containing protein	0	4	0	4
Acin1	apoptotic chromatin condensation inducer in the	0	4	0	4
Ddx18	ATP-dependent RNA helicase DDX18	0	4	0	4
Fads2	fatty acid desaturase 2	0	4	0	4
Flnc	filamin-C	0	4	0	4
Krt33a	keratin, type I cuticular Ha3-I	0	4	0	4
Lama3	laminin subunit alpha-3	0	4	0	4
LOC100		0	4	0	4
Nnt	NAD(P) transhydrogenase, mitochondrial	0	4	0	4
Nvl	nuclear valosin-containing protein-like	0	4	0	4
Prpf6	pre-mRNA-processing factor 6	0	4	0	4
Rpl22	60S ribosomal protein L22	0	4	0	4
Rpl23	60S ribosomal protein L23	0	4	0	4
Tln1	talin-1	0	4	0	4
Wdr36	WD repeat domain 36	0	4	0	4

D1Pas1	putative ATP-dependent RNA helicase PI10		4	0	4
Fscn1	fascin	2	6	1	4
Hnrnph	heterogeneous nuclear ribonucleoprotein H	2	6	1	4
Hspb1	heat shock protein beta-1	0	4	1	4
Actn4	alpha-actinin-4	0	4	2	4
Dst	dystonin	0	4	2	4
Gemin5	gem-associated protein 5	0	4	2	4
Hspa4	heat shock 70 kDa protein 4	0	4	2	4
Odf2l	outer dense fiber protein 2-like	0	4	2	4
Bicd1	protein bicaudal D homolog 1	0	4	3	4
Sarnp	SAP domain-containing ribonucleoprotein	0	4	3	4
Vps35	vacuolar protein sorting-associated protein 35	4	8	4	4
Gucy2c	heat-stable enterotoxin receptor	2	6	4	4
Krt35	keratin, type I cuticular Ha5	2	6	4	4
Plk2	serine/threonine-protein kinase PLK2	2	6	6	4
241000	UPF0533 protein C5orf44 homolog	1	4	0	3
Fam98c	uncharacterized protein LOC73833	1	4	0	3
Rpl28-		1	4	0	3
Shisa9	protein shisa-9	1	4	0	3
Abcc2	canalicular multispecific organic anion transporter 1	0	3	0	3
Alyref	THO complex subunit 4	0	3	0	3
Cacna2	voltage-dependent calcium channel subunit alpha-	0	3	0	3
Dsg1a	desmoglein-1-alpha	0	3	0	3
E2f2	transcription factor E2F2	0	3	0	3
Eed	polycomb protein EED	0	3	0	3
Eftud2	116 kDa U5 small nuclear ribonucleoprotein	0	3	0	3
Fam65b	protein FAM65B	0	3	0	3
Grpel1	grpE protein homolog 1, mitochondrial	0	3	0	3
Hdac2	histone deacetylase 2	0	3	0	3
Kif1b	kinesin-like protein KIF1B	0	3	0	3
Lgals3b	galectin-3-binding protein	0	3	0	3
Nup160	nuclear pore complex protein Nup160	0	3	0	3
P2rx3	P2X purinoceptor 3	0	3	0	3
Pcnx13	pecanex-like protein 3	0	3	0	3
Pml	protein PML	0	3	0	3
Smc3	structural maintenance of chromosomes protein 3	0	3	0	3
Smg6	telomerase-binding protein EST1A	0	3	0	3
Tnr	tenascin-R	0	3	0	3
Vezt	vezatin	0	3	0	3
Vps16	vacuolar protein sorting-associated protein 16 homolog	0	3	0	3
Wdr46	WD repeat-containing protein 46	0	3	0	3
Yes1	tyrosine-protein kinase Yes	0	3	0	3
Arid5a	AT-rich interactive domain-containing protein 5A	0	3	1	3
Ddx58	probable ATP-dependent RNA helicase DDX58	0	3	1	3
Gdf15	growth/differentiation factor 15	0	3	1	3
Hinfp	histone H4 transcription factor	0	3	1	3
Sf3b3	splicing factor 3B subunit 3	0	3	1	3
Slc25a1	calcium-binding mitochondrial carrier protein Aralar2	0	3	1	3
Stag2	cohesin subunit SA-2	0	3	1	3
Taf1	transcription initiation factor TFIID subunit 1	0	3	1	3
Lrpprc	leucine-rich PPR motif-containing protein,	1	4	2	3
Herc2	E3 ubiquitin-protein ligase HERC2	0	3	2	3
Rere	arginine-glutamic acid dipeptide repeats protein	0	3	2	3
Hnrnpa	heterogeneous nuclear ribonucleoprotein A/B	4	7	3	3
Fras1	extracellular matrix protein FRAS1	2	5	3	3
Hnrnpd	heterogeneous nuclear ribonucleoprotein D0	2	5	3	3
Chrna1	neuronal acetylcholine receptor subunit alpha-10	0	3	3	3
Dock9	dedicator of cytokinesis protein 9	0	3	3	3
Erc1	ELKS/Rab6-interacting/CAST family member 1	0	3	3	3
Hnrnpa	heterogeneous nuclear ribonucleoprotein A0	0	3	3	3

Pard6g	partitioning defective 6 homolog gamma	0	3	3	3
Prpf8	pre-mRNA-processing-splicing factor 8	0	3	3	3
Ube2t	ubiquitin-conjugating enzyme E2 T	0	3	3	3
Vash1	vasohibin-1	1	4	4	3
Mei4	UPF0623 protein	0	3	4	3

Table S2. Primer sequences.

cloning primer		
Suv39h2	f	ggggacaagttgtacaaaaaagcaggcttaactatggcggcgccaggggcc
	r	ggggaccactttgtacaagaaagctgggtcggtgaggtaacctctgcaa
M12,M13+M14 pGex6P1		
Suv4-20h2 pGex6P1: 350f	f	GGGATCCgtcctccgcactgcctgtgt
Suv4-20h2 pGex6P1: 385f	f	GGGATCCcgctggaccccacaacag
Suv4-20h2 pGex6P1: 412r	r	GGAATTCctaggctaggcgggtaagttc
Suv4-20h2 pGex6P1: 385r	r	GGAATTCctagcgagctctggggcgcgag
M7	f	ggggacaagttgtacaaaaaagcaggcttaacctgacccgcctagccccagc
	r	ggggaccactttgtacaagaaagctgggtcggctcaccactattgat
M12	f	ggggacaagttgtacaaaaaagcaggcttaacctggtcctccgcactgcctgt
	r	ggggaccactttgtacaagaaagctgggtcggctaggcgggtaagttc
M13	f	ggggacaagttgtacaaaaaagcaggcttaacctggtcctccgcactgcctgt
	r	ggggaccactttgtacaagaaagctgggtcgcgagctctggggcgcgag
M14	f	ggggacaagttgtacaaaaaagcaggcttaacctgcgctggaccccacaaca
	r	ggggaccactttgtacaagaaagctgggtcggctaggcgggtaagttc
delta M12	f	attgcctcctccagccctgccagctgct
	r	gctggaggaggcaatgaagcctggc
HP1 beta	f	ggggacaagttgtacaaaaaagcaggcttaactatggggaaaaagcaaaac
	r	ggggaccactttgtacaagaaagctgggtcattcttgcctctttttgtc
HP1 gamma	f	ggggacaagttgtacaaaaaagcaggcttaaaaaatggcctccaataaaactac
	r	ggggaccactttgtacaagaaagctgggtccttgcctcatctcaggac
HP1 alpha	f	ggggacaagttgtacaaaaaagcaggcttaagacatgggaaagaagacc
	r	ggggaccactttgtacaagaaagctgggtcgcctcttcgcgcttcttttc
Suv4-20h1 10	f	GGGGACAAGTTTGTACAAAAAAGCAGGCTACaactatggtggt
Suv4-20h1 883	r	GGGGACCACTTTGTACAAGAAAGCTGGGTCTcgcgttcagctct
Suv4-20h1 392	r	GGGGACCACTTTGTACAAGAAAGCTGGGTCTtacttgcattgt
Suv4-20h1 392	f	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAacctatgtcta
Suv4-20h1 847	r	GGGGACCACTTTGTACAAGAAAGCTGGGTCTgtcatcatcgaa
Suv4-20h1 496	r	GGGGACCACTTTGTACAAGAAAGCTGGGTCTctccctttcttctt
ChIP primer		
Major satellites	f	GACGACTTGAAAAATGACGAAATC
	r	CATATTCCAGGTCCTTCAGTGTGC
Chr. 6	f	GATGGGAAAGCGTTGTTAGC
chr6:90702961+9070304	r	AAGGACAGCTCCTTTTTTCAGG
Chr. 7	f	AATTTCACTGCGATCCTTGC
chr7:45669448+4566954	r	GCCGGAATTATGGCTCTATG
Chr. 10	f	TTCTGCTAAAGCCTGGACTTG
chr10:76574384+765745	r	TGCTTGCAAGTGGCTAAGG
Chr. 5	f	ATGCTCACGTCCTTGTCCAGAA
chr5:147617294+147617	r	TCTGGCAGCCTTCAACGTTTGT
Chr. 6-1	f	TGCAGGTGGGATTAAGTGTG
chr6:122673061+122673	r	CTACCCACCCCTATTCTCC
Chr10-1	f	TGGGTGCCGTATGCCACATTAT
chr10:79154346+791544	r	TTTCTGGCCATCCGCACCTTAT
genotyping primer for		
MGC tag 1f	f	CACTCCCACTGTCCTTTCCTAA
MGC tag 1r	r	CAAACCTCACAGAGGCCACCTA
MGC tag 2f	f	TCGCATTGTCTGAGTAGGTGTC

2.6 CENP-C facilitates the recruitment of M18BP1 to centromeric chromatin.

Article: Dambacher et al. 2012, Nucleus

CENP-C facilitates the recruitment of M18BP1 to centromeric chromatin

Silvia Dambacher,^{1,2,†} Wen Deng,^{1,3,†} Matthias Hahn,^{1,2,†} Dennis Sadic,^{1,2} Jonathan J. Fröhlich,^{1,2} Alexander Nuber,^{1,2} Christian Hoischen,⁴ Stephan Diekmann,⁴ Heinrich Leonhardt^{1,3} and Gunnar Schotta^{1,2,*}

¹Ludwig Maximilians University and Munich Center for Integrated Protein Science (CiPS^M); Munich, Germany; ²Adolf-Butenandt-Institute; Munich, Germany;

³Department of Biology II; Ludwig Maximilians University; Munich, Germany; ⁴Leibniz Institute for Age Research; Fritz Lipmann Institute; Jena, Germany

[†]These authors contributed equally to this paper.

Centromeres are important structural constituents of chromosomes that ensure proper chromosome segregation during mitosis by providing defined sites for kinetochore attachment. In higher eukaryotes, centromeres have no specific DNA sequence and thus, they are rather determined through epigenetic mechanisms. A fundamental process in centromere establishment is the incorporation of the histone variant CENP-A into centromeric chromatin, which provides a binding platform for the other centromeric proteins. The Mis18 complex, and, in particular, its member M18BP1 was shown to be essential for both incorporation and maintenance of CENP-A.

Here we show that M18BP1 displays a cell cycle-regulated association with centromeric chromatin in mouse embryonic stem cells. M18BP1 is highly enriched at centromeric regions from late anaphase through to G1 phase. An interaction screen against 16 core centromeric proteins revealed a novel interaction of M18BP1 with CENP-C. We mapped the interaction domain in M18BP1 to a central region containing a conserved SANT domain and in CENP-C to the C-terminus. Knock-down of CENP-C leads to reduced M18BP1 association and lower CENP-A levels at centromeres, suggesting that CENP-C works as an important factor for centromeric M18BP1 recruitment and thus for maintaining centromeric CENP-A.

Introduction

Centromeres are sites for kinetochore attachment during mitosis. In order to prevent chromosome segregation defects, cells have to ensure that each chromosome has one functional centromere. Centromeres have no fixed DNA sequence that can be recognized by specific binding proteins, therefore it is assumed that epigenetic mechanisms ensure maintenance of the centromeric structure. The histone H3 variant CENP-A is a central component of centromeric chromatin. CENP-A aids in recruiting numerous proteins that build the constitutive centromere-associated network (CCAN),^{1–3} an essential step in establishing a proper kinetochore structure.⁴ Two proteins directly bind CENP-A and have the potential to bridge centromeric chromatin with kinetochore components. The first protein, CENP-C, recognizes the C-terminal region of CENP-A through an internal region.⁵ The C-terminus of CENP-C mediates its dimerization,^{6,7} the extreme N-terminus interacts with the Mis12 complex, which, in turn, bridges to outer kinetochore components.⁸ The second protein directly recognizing CENP-A is CENP-N,⁹ which also interacts with other centromeric components. Notably, disruption of either CENP-C or CENP-N leads to reduced levels of CENP-A at centromeres, suggesting that both proteins have additional functions in establishment or maintenance of centromere identity.^{5,9}

The incorporation of CENP-A into the centromere is a strictly cell cycle-regulated process. During replication of centromeres, CENP-A is equally distributed onto the daughter strands, diluting the amount per centromere to 50%. To preserve centromere function, CENP-A needs to be subsequently replenished. Expression levels of CENP-A peak in G₂ phase, though incorporation into the centromere only occurs in late mitosis and early G₁ phase.^{10–13} The histone chaperone that mediates incorporation of CENP-A is the Holliday junction-recognizing protein (HJURP).^{14,15} HJURP can incorporate CENP-A only in domains that show a signature of actively transcribed chromatin.¹⁶ Therefore, centromeric chromatin needs to be prepared (licensed), by currently unknown mechanisms. Mis18 α , Mis18 β and M18BP1 which form the Mis18 complex in humans have been suggested to play an important role in this licensing mechanism.^{17–19} Disruption of Mis18 complex components leads to failure in CENP-A incorporation,^{17,19} which could be explained by lack of HJURP recruitment to centromeres.^{20,21} Neither of the Mis18 complex proteins directly interact with CENP-A,⁹ therefore, an important question within the understanding of CENP-A establishment is how this complex is specifically targeted to centromeric chromatin.

Here we show that M18BP1 is a cell cycle-regulated component of centromeric chromatin. By screening 16 CCAN proteins we

*Correspondence to: Gunnar Schotta; Email: gunnar.schotta@med.uni-muenchen.de
Submitted: 10/29/11; Revised: 12/04/11; Accepted: 12/06/11
<http://dx.doi.org/10.4161/nucl.3.1.18955>

identify CENP-C as a novel interaction partner of M18BP1. We mapped the interaction domain to a central region of M18BP1 encompassing the conserved SANT domain. CENP-C facilitates the recruitment of M18BP1 to centromeric chromatin during specific stages of the cell cycle, as RNAi depletion of CENP-C leads to reduced levels of centromeric M18BP1. In summary, our work identifies CENP-C as an important centromere component that recruits M18BP1 to centromeric chromatin.

Results

M18BP1 is a centromere-associated protein in mouse ES (mES) cells. In HeLa cells, prominent centromeric association of Mis 18 complex members was observed from late mitosis (ana/telophase) until the end of G₁ phase.¹⁷ In order to determine the localization of M18BP1 in mouse cells we generated M18BP1 knock-in mES cells (K1B2) by introducing an EGFP tag into the endogenous M18BP1 locus (Fig. S1A). The K1B2 cells express M18BP1 at near endogenous levels (Fig. S1B), suggesting that the transcriptional

regulation of M18BP1 is not impaired by the alterations to this locus. We determined the localization of M18BP1-EGFP in these cells by comparing the EGFP signal with CENP-A staining to ask whether M18BP1 localizes to the centromeres. Three classes of staining patterns could be observed (Fig. 1A): (1) diffuse nuclear; (2) weak centromeric; and (3) strong centromeric foci. We then asked which cell cycle stage would correspond to the strong centromeric association of M18BP1. K1B2 cells show the typical cell cycle profile of mES cells, that is, the majority of cells are in S phase with an additional high percentage of cells in G₂/M phase. In contrast to other frequently analyzed cell types, such as HeLa cells and mouse fibroblasts, the population of G₁ cells is comparably low in mES cells. We visualized the different cell cycle stages in K1B2 cells using specific markers. First we expressed RFP-tagged PCNA in K1B2 cells, which is an indicator of different S phase stages.²² All K1B2 cells which showed defined PCNA dots, indicative of ongoing replication showed weak or diffuse M18BP1-EGFP signals (Fig. 1B). A significant number of cells outside S phase showed comparatively stronger signals, suggesting that centromeric

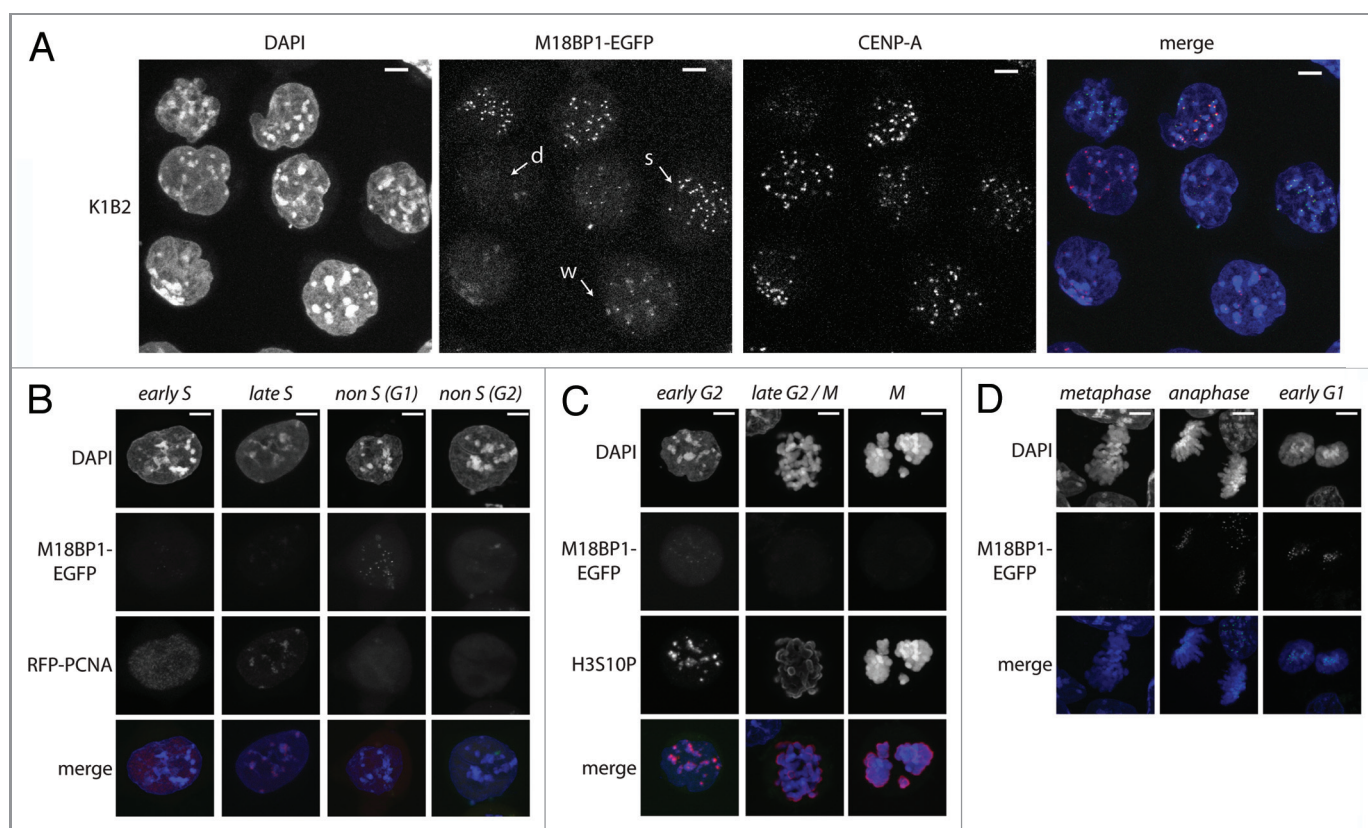


Figure 1. M18BP1 associates with centromeres in a cell-cycle dependent manner. (A) Localization of endogenously tagged M18BP1-EGFP in the K1B2 mES cell line. K1B2 cells were stained for CENP-A and confocal stacks were recorded. Maximum intensity projections are shown. M18BP1-EGFP showed different patterns: strong enrichment at centromeres (s), weak enrichment (w), no enrichment/diffuse nuclear (d). Scale bars are 20 μm. (B) M18BP1 distribution during S phase. K1B2 cells were transfected with a RFP-PCNA expression construct to detect cells in different S phase stages. M18BP1-EGFP showed intermediate to low centromeric enrichment throughout S phase. Cells which are not in S phase fall into two different staining patterns: M18BP1 is highly enriched at centromeres (presumably G1) and cells with low/no centromeric M18BP1 signals (presumably G2). (C) M18BP1 distribution in G2/M phase. K1B2 cells were stained with H3S10P antibodies to visualize different stages of G2 and M phase. Starting from early G2 phase (weak H3S10P signal) to M phase (strong H3S10P signal) M18BP1 appeared to be largely absent from centromeres. (D) M18BP1 localization in different mitotic stages. In metaphase cells, M18BP1 is absent from centromeres, however, starting from late anaphase, M18BP1 showed strong signals at centromeric regions. Scale bars in (B–D) are 5 μm.

M18BP1 association is low throughout S phase (Fig. 1B) and is only enriched in G₁ or G₂/M phase. We therefore tested whether M18BP1 begins to be enriched after S phase, in G₂/M. In order to visualize G₂/M cells, we performed immunofluorescence staining for H3S10 phosphorylation.²³ All K1B2 cells that were positive for H3S10P showed only weak centromeric M18BP1 signals (Fig. 1C, early G₂), leading to the conclusion that the highly enriched M18BP1 signals appear in G₁ phase cells. Interestingly, in G₂ phase cells, M18BP1 does not seem to associate with all centromeres as we detect numerous CENP-A spots without M18BP1 enrichment (Fig. S2). Late G₂ and prometaphase cells did not show significant centromeric signals for M18BP1 (Fig. 1C, late G₂/M). Finally, we tested at which step after mitosis M18BP1 starts being localized to centromeric regions by examining distinct mitotic stages in K1B2 cells. Notably, we find that metaphase cells still display very low M18BP1 signals, though as soon as cells enter anaphase/telophase, M18BP1 is highly enriched at centromeres (Fig. 1D). In those cells centromeric CENP-A signals are still relatively low as deposition of new CENP-A only occurs at later stages in the cell cycle.¹² In summary, our localization analysis demonstrates that in mES cells, M18BP1 is not constitutively enriched at centromeric chromatin but rather associates with centromeres from anaphase continuing to G₁ phase. This is the time when CENP-A incorporation takes place.

Centromere interaction screen for M18BP1. M18BP1 is known to be important for preparing centromeric chromatin for CENP-A incorporation. However, still very little is known about how M18BP1 actually recognizes centromeric chromatin. We

pursued the idea that M18BP1 might be recruited through interaction with components of the CCAN network. To test this hypothesis we performed an F3H interaction screen of M18BP1 with proteins of the CCAN network. The F3H interaction assay utilizes a BHK cell line with a lac operator repeat array stably integrated into its genome (Fig. 2A). This cell line was transfected with an expression vector encoding the lac repressor (lacI) which directly binds to the lac operator sequence fused with a GFP binding protein (GBP). These cells further expressed the EGFP tagged bait protein (M18BP1-EGFP) and individual RFP/mCherry tagged prey proteins (CCAN proteins). M18BP1-EGFP is bound by the lacI-GBP fusion protein at the lac operator arrays and can be detected at the well-discernible nuclear lacO focus. Prey protein interaction with M18BP1 is identified by localization to this nuclear focus (Fig. 2A). The red/green signal intensity ratio provides a measure for the strength of the tested interaction.

We tested M18BP1 for interaction with 16 proteins of the CCAN network using the F3H assay. In agreement with previous analyses we did not detect a direct interaction with CENP-A (Fig. 2B). Interestingly, we found a strong interaction with another protein of the inner centromere, CENP-C. None of the other CCAN proteins that we tested showed significant interaction with M18BP1 (Fig. 2C).

M18BP1 harbors two evolutionarily conserved domains, the SANT domain in the C-terminal part of the protein and the SANTA domain which is toward the N-terminus. To test which region of M18BP1 participates in the interaction with CENP-C

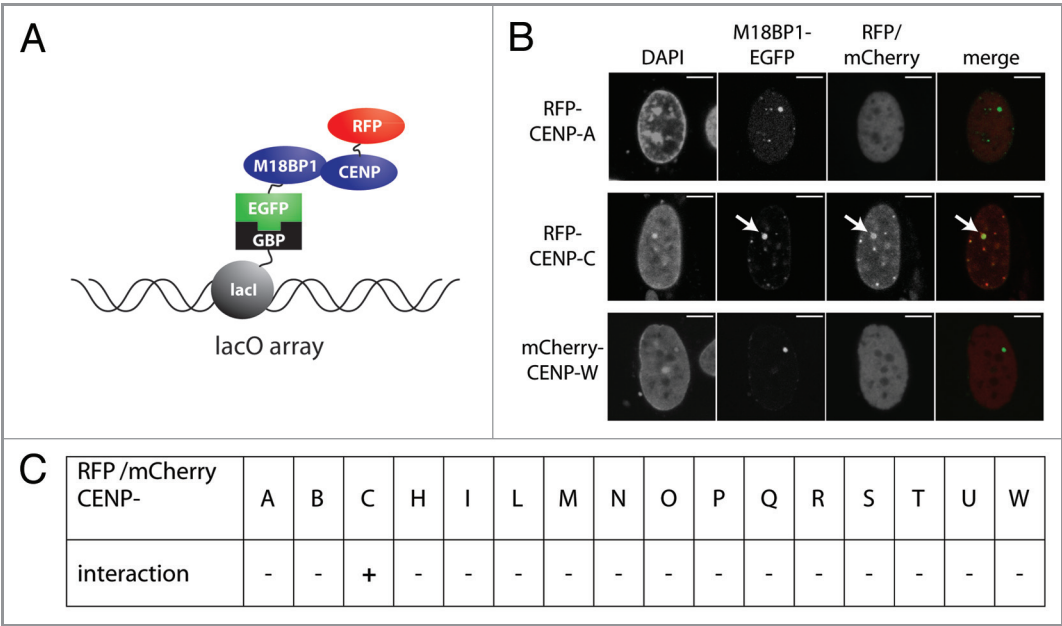


Figure 2. F3H interaction screen for M18BP1 interaction partners. (A) Scheme depicting the F3H screening strategy. Cells containing a lac operator array were transfected with plasmids expressing a lac repressor-GBP fusion protein, M18BP1-EGFP and mCherry/RFP-CCAN proteins. The lac repressor binds to the lac operator array and through the GBP recruits M18BP1-EGFP. CCAN proteins interacting with M18BP1 are consequently enriched at the lac operator array. (B) Representative examples for M18BP1 interacting (CENP-C) and non-interacting (CENP-A and CENP-W) proteins are shown. Scale bar is 5µm. (C) Summary of interaction tests between M18BP1 and CCAN proteins. Interactions were tested with the F3H assay using M18BP1-EGFP and 16 RFP or mCherry fusions with CCAN proteins. From all 16 tested CCAN proteins, only CENP-C showed a clear interaction with M18BP1.

we tested a panel of M18BP1 truncated proteins (Fig. 3A). The N-terminus of M18BP1 (M1, aa1–440) showed no interaction with CENP-C, but the C-terminus (M2, aa441–998) clearly interacted (Fig. 3B). The central region (M3, aa325–800) and a truncated protein lacking the SANTA domain (M4, aa441–800) displayed clear interactions with CENP-C. In order to then test whether the SANT domain is sufficient for the CENP-C interaction, we assessed a truncated protein harboring only the SANT domain (M5, aa735–800). This failed to interact with CENP-C, suggesting that additional parts of M18BP1 participate in this interaction. We scrutinized these observations by quantifying the interactions in several hundred cells per construct through measurement of the ratio between red and green intensity values at the M18BP1-EGFP foci. Although, by confocal imaging we can detect red/green colocalization of M18BP1-M4 and CENP-C in 73% of the cells (Fig. 3B), the average red/green signal ratio is relatively low (Fig. 3C). This, however, can be explained by relatively low expression levels of RFP-CENP-C in the combination with M18BP1-M4. In summary, our F3H data show that CENP-C interacts with a central region of M18BP1 comprising the SANT domain.

M18BP1 directly interacts with CENP-C. We then aimed to further define the M18BP1-CENP-C interaction. First, we wanted to analyze whether M18BP1 co-localizes with CENP-C. To do this, we transfected K1B2 cells with a plasmid expressing RFP-tagged CENP-C and performed confocal imaging. We found many cells showing a clear overlap between M18BP1-EGFP and RFP-CENP-C signals. However, there was also a large percentage of cells with prominent CENP-C signals, with no M18BP1 co-localization (Fig. 4A). These data suggest that the interaction between these two proteins is highly regulated in vivo.

In order to test whether CENP-C and M18BP1 can interact in vivo, we performed co-immunoprecipitation experiments. We transfected HEK293 cells with expression plasmids for CENP-C-EGFP and myc-M18BP1, prepared nuclear extract and purified CENP-C-EGFP using GFP trap affinity beads. In the bound material we could clearly detect co-purification of the EGFP-tagged CENP-C and the myc-tagged M18BP1 (Fig. 4B). We then wanted to further map the interaction domains between M18BP1 and CENP-C using in vitro binding assays. CENP-C has several conserved domains which have already been implicated in different biochemical interactions and in vivo functions, such as

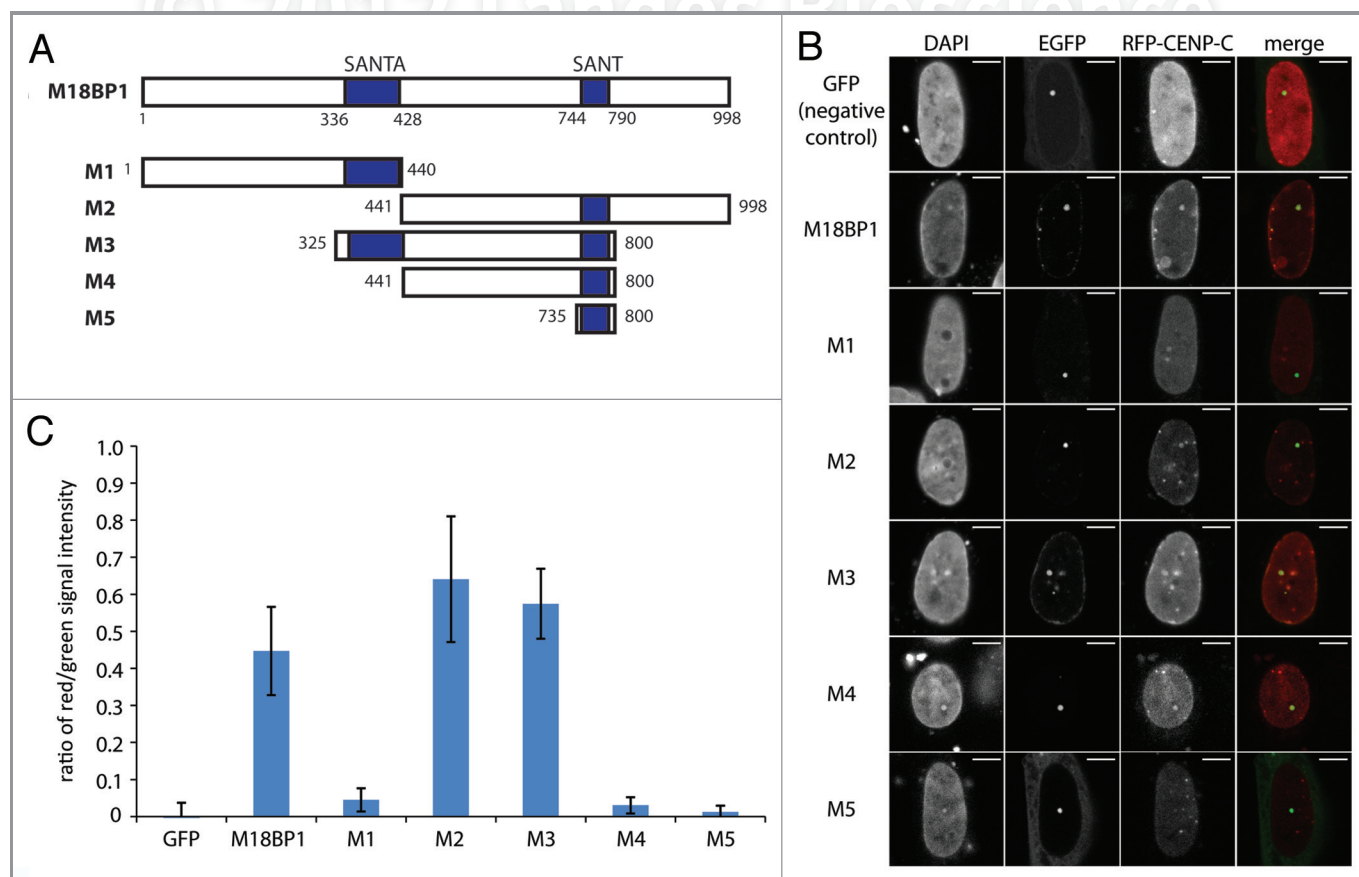


Figure 3. Mapping of the M18BP1-CENP-C interaction domains by F3H. (A) Scheme of M18BP1 truncations used in the interaction tests. (B) Representative F3H images of the negative control (GFP only) and EGFP tagged M18BP1 truncations tested with RFP-CENP-C. Overlap of red and green signals at the nuclear *lacO* focus indicates interaction. (C) Quantification of the F3H M18BP1-CENP-C interaction data. Intensities of red and green signals at the nuclear *lacO* focus were measured in several hundred cells each. The ratio between red and green signals was determined to measure the strength of the tested interactions.

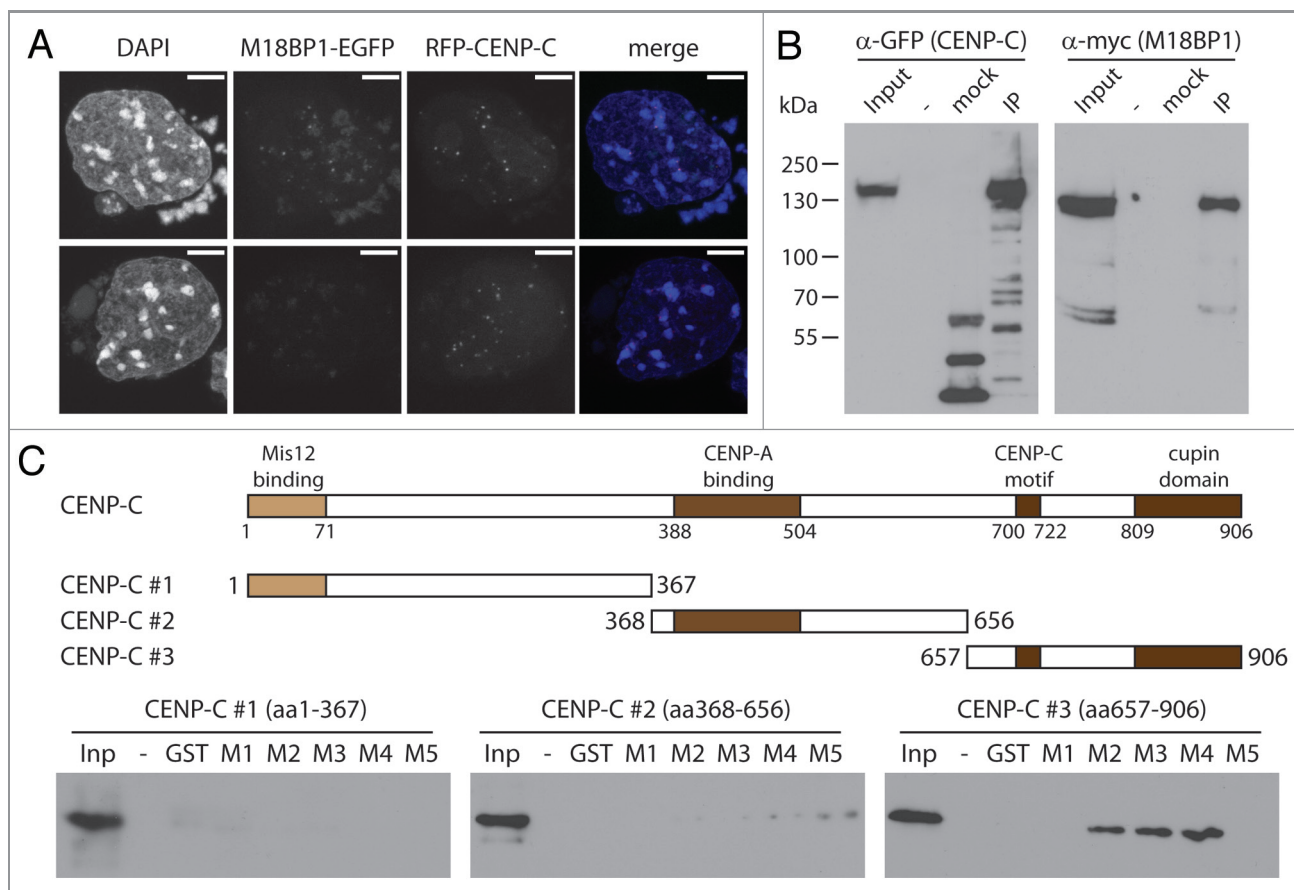


Figure 4. M18BP1 and CENP-C interact in vitro. (A) Co-localization of RFP-CENP-C and M18BP1-EGFP. K1B2 cells were transfected with a RFP-CENP-C expression construct. Maximum intensity projections of two representative staining patterns are shown. Scale bar is 5 μ m. (B) Co-immunoprecipitation of CENP-C and M18BP1. HEK293FT cells were transfected with expression plasmids for EGFP-CENP-C and myc-M18BP1. Nuclear extracts from these cells were incubated with agarose beads (control) and GFP-Trap affinity beads to enrich for EGFP-CENP-C and interacting bound proteins. Protein gel blot analysis shows the nuclear extract (Inp), proteins bound to agarose beads (mock) and proteins that were enriched with GFP-Trap agarose beads (IP). An empty lane is indicated by “-”. EGFP-CENP-C and myc-M18BP1 were detected using antibodies against GFP and myc, respectively. (C) Interaction tests between M18BP1 and CENP-C truncations. The scheme shows the domain structure of mouse CENP-C and the truncation constructs that were used in this assay. Recombinant GST-tagged M18BP1 truncations (M1-M5) were incubated with in vitro translated myc-CENP-C truncation proteins and bound to GST beads. The bound CENP-C protein truncations were detected using myc antibody. Only the C-terminal CENP-C fragment showed clear interaction with M18BP1. The M18BP1 fragments M1-M5 are depicted in **Figure 3A**.

connection to the outer kinetochore proteins, CENP-A binding and CENP-C dimerization (**Fig. 4C**, schematic). We generated in vitro translated proteins of three CENP-C truncations and tested their interaction with recombinant GST tagged M18BP1 truncations (M1-M5). In these assays we could only detect significant interaction of M18BP1 with CENP-C #3, containing the CENP-C motif and the cupin domain (**Fig. 4C**). These data provide an extension of the F3H analysis; confirming that the large central region of M18BP1 is required for CENP-C binding.

CENP-C is required for the recruitment of M18BP1 to centromeres. Our data show that M18BP1 interacts with CENP-C in vitro and in vivo. CENP-C itself binds to centromeres through direct interaction with CENP-A. We therefore hypothesized that CENP-C facilitates the recruitment of M18BP1 to centromeric chromatin. In order to test this hypothesis we performed CENP-C knock-down experiments in K1B2 cells in which we could easily assess the localization of endogenously

expressed M18BP1-EGFP. We prepared pLKO-based lentiviral vectors with three independent shRNA oligos against CENP-C and one control oligo containing an unrelated sequence (**Table S1**). CENP-C knock-down cells were analyzed by qPCR five days post infection to determine the knock-down efficiency of the individual oligos. Importantly, all three knock-down oligos resulted in effective downregulation of CENP-C mRNA (**Fig. 5A**) and protein (**Fig. S3**), with shCENP-C #3 showing the strongest knock-down. Crucially, the expression level of M18BP1 was unchanged. CENP-C knock-down did not lead to significant changes in the cell cycle profile of K1B2 cells, however, we did notice an increase in the number of cells with sub-G1 DNA content (**Fig. S4**) as well as reduced cell numbers at day five after knock-down (**Fig. S5**), indicating elevated cell death upon CENP-C knock-down.

To test whether CENP-C affects the localization of M18BP1 we investigated M18BP1-EGFP and CENP-A patterns in the CENP-C knock-down cells. In control knock-down cells, M18BP1-EGFP

shows the typical distribution of different centromere enrichment levels: strong, weak and diffuse nuclear (Fig. 5B, arrows). We observed that the number of cells with centromeric association of M18BP1 was reduced in all three knock-down cell lines (Fig. 5B). In order to quantify this phenotype we determined the distribution of M18BP1 staining patterns in control and CENP-C knock-down cells. Importantly, in all knock-down cell lines the percentage of cells with “weak” M18BP1-EGFP enrichment at centromeres was reduced, whereas M18BP1-EGFP “diffuse” cells were increased (Fig. 5C). In knock-down shCENP-C #3 we even detected reduced numbers of M18BP1 ‘strong’ cells, suggesting that more efficient CENP-C knock-down more severely impairs centromeric M18BP1 recruitment. We then asked whether the reduced centromeric M18BP1 recruitment corresponds to specific cell cycle stages by co-staining of control and CENP-C knock-down cells with specific cell cycle markers (Fig. S6). In shControl cells we could reproduce the results of our initial cell cycle analysis in K1B2 cells: G₁ cells showed strong centromeric signals, G₂ cells showed weak signals. Importantly, upon CENP-C knock-down we detected G₁ and G₂ phase cells which had clearly lost centromeric M18BP1 (Fig. S6),

indicating that the role of CENP-C in ensuring centromeric M18BP1 localization is not restricted to a particular cell cycle stage.

M18BP1 was proposed to ‘prime’ centromeres for deposition of CENP-A.^{17,19} In order to investigate whether reduced centromeric M18BP1 recruitment would also lead to less efficient CENP-A incorporation, we divided CENP-A staining patterns into low, medium and high and determined the percentage of cells showing these patterns in control and CENP-C knock-down cells. In particular knock-down shCENP-C #3 which had the strongest effect on centromeric M18BP1 recruitment lead to significantly reduced CENP-A levels (Fig. 5C).

In summary our data demonstrate that the interaction between CENP-C and M18BP1 is an important recruitment mechanism for M18BP1 to centromeric chromatin, which appears necessary for the correct deposition of CENP-A.

Discussion

The deposition of CENP-A into centromeric chromatin is essential to ensure proper segregation of chromosomes. The

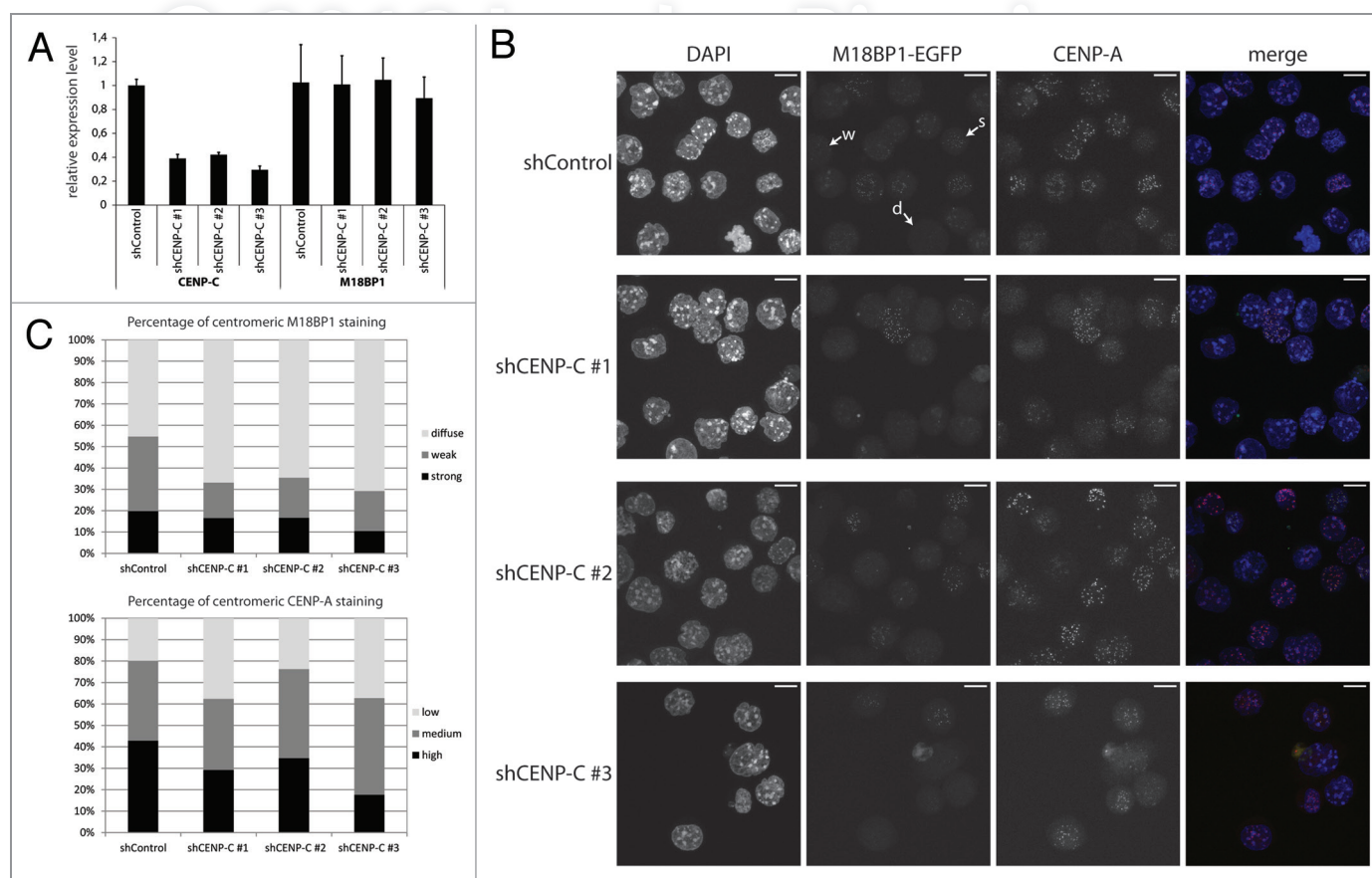


Figure 5. CENP-C knock-down leads to impaired centromeric recruitment of M18BP1. (A) RT-qPCR for CENP-C and M18BP1 five days after knock-down. Expression levels in the control knock-down cell line (shControl) and the three CENP-C knock-down cell lines (shCENP-C #1-#3) were normalized to the geometric mean of GAPDH and Actin. (B) Representative maximum intensity projections of confocal stacks of control and CENP-C knock-down cells that were stained for CENP-A. Arrows point to example cells for the three classes of M18BP1 signals: strong (s), weak (w) and no enrichment/diffuse nuclear (d). Scale bars are 10µm. (C) Quantification of the M18BP1 signals in control vs. CENP-C knock-down cell lines. M18BP1 and CENP-A staining patterns were classified in several hundred cells. The bar graph depicts the percentages of each class.

Mis18 complex member M18BP1 was shown as an essential factor to prepare centromeric chromatin for CENP-A deposition and to ensure its maintenance.^{17,19,24} Our data constitute the first analysis of endogenously expressed M18BP1 in mES cells. We report that M18BP1 associates with centromeric chromatin during distinct cell cycle stages. M18BP1 shows highest abundance at centromeres from anaphase through to late G₁ phase. These data are in agreement with observations in human cells, where M18BP1 also associates with centromeres starting from late telophase through to G₁ phase¹⁷ and suggest that M18BP1-mediated processes might be evolutionarily conserved in higher mammals.

We have furthered the understanding of how M18BP1 is recruited to centromeres through identification of a novel interaction between the C-terminus of CENP-C with a central region in M18BP1, which contains a SANT domain. The SANT domain is highly conserved and found in many chromatin-associated proteins, but very little is known about its potential functions. It has been implicated in the mediation of protein-protein interactions and binding to histone modifications.²⁵ We do not detect a direct interaction between the isolated M18BP1 SANT domain and CENP-C, however, it is possible that this domain is only functional in a larger protein context. More detailed experiments are necessary to further understand the functional roles of this domain in M18BP1. Our interaction data are consistent with a very recent study which appeared during the preparation of this manuscript.²⁶ Moree et al. found that *X. laevis* xM18BP1 isoforms interact with xCENP-C, and they could also show that human M18BP1 interacts with human CENP-C. In their study the interaction domain with xM18BP1 was mapped to the C-terminus of xCENP-C containing the CENP-C motif and the cupin domain. Mouse CENP-C (906aa) is much smaller than xCENP-C (1400aa), however, the major domains, such as CENP-A binding domain, CENP-C motif and cupin domain are conserved. We could show that in the mouse the interaction with M18BP1 is also mediated through a C-terminal fragment of CENP-C containing the CENP-C motif and the cupin domain, and thus the M18BP1 interaction site in CENP-C seems to be evolutionarily conserved.

Our data moreover demonstrate an important function for CENP-C in mediating the centromeric recruitment of M18BP1. When CENP-C protein levels are reduced to around 40–50% (shCENP-C #1 and #2) we found reduced numbers of “weak” centromeric M18BP1 cells. The numbers of “strong” centromeric M18BP1 cells, a staining pattern which we found characteristic for G₁ phase cells, seemed to be unaltered. More severely reduced CENP-C levels (shCENP-C #3) resulted in lower numbers of cells with “weak” and “strong” centromeric M18BP1. Our cell cycle marker analysis revealed that the role of CENP-C in mediating centromeric M18BP1 recruitment is not restricted to selective cell cycle stages. However, we cannot exclude the possibility that CENP-C has additional functions that indirectly regulate the centromeric association of M18BP1 during distinct cell cycle phases. In this context it is interesting to note that when we transiently express CENP-C in K1B2 cells, we find a high percentage of cells in which CENP-C is abundantly

associated with centromeres, but M18BP1 does not co-localize. It is therefore plausible to assume that further regulatory mechanisms exist, e.g., post-translational modifications, which influence the in vivo interaction between these two proteins. Both M18BP1 and CENP-C can be phosphorylated and sumoylated at multiple sites.^{27–30} It will be challenging to understand how these modifications are regulated and how they influence interactions between the different centromeric proteins in a cell cycle-dependent manner.

The centromeric recruitment of M18BP1 appears important for correct deposition of CENP-A. In particular strong depletion of CENP-C with knock-down oligo shCENP-C #3 leads to reduced levels of centromeric CENP-A. These data are consistent with Moree et al., which demonstrate in the *Xenopus* system that upon xCENP-C depletion, centromeric deposition of new CENP-A is impaired.²⁶ The failure to correctly establish CENP-A might be due to loss of centromeric M18BP1 at critical cell cycle stages. In human cells, CENP-A deposition is mediated by HJURP during G₁ phase. Loss of M18BP1 leads to reduced HJURP association with centromeres and consequently to reduced deposition of newly synthesized CENP-A.^{14,15} In our mES cell system strong CENP-C knock-down results in cells which lose M18BP1 during G₁ phase when CENP-A deposition normally occurs. We therefore postulate that those cells will also have problems in correctly establishing centromeric CENP-A patterns. We do not detect a large number of cells which have lost CENP-A upon CENP-C knock-down. We think that this could be explained by the high cell lethality of the CENP-C knock-down. Therefore, at the current stage of analysis, we cannot distinguish whether critically low CENP-A levels would induce apoptosis in ES cells, or whether CENP-C has additional functions that could be critical for survival of ES cells. Also, the functions of M18BP1 need to be investigated in more detail to understand how the centromeric recruitment of this molecule drives the subsequent deposition of CENP-A during G₁ phase and whether M18BP1 features additional roles during other cell cycle stages when its centromeric recruitment is much lower but still detectable.

Materials and Methods

M18BP1 knock-in cell line. The M18BP1-EGFP targeting constructs were obtained using the recombineering cloning technique described previously.³¹ To generate retrieval and mini-targeting vectors, PCR fragments were amplified from the BAC clone RP23–396P24 (Children's Hospital Oakland Research Institute). For the retrieval plasmid, PCR fragments were cloned into the pL253 plasmid using NotI, HindIII and SpeI. A genomic region of 7 kb, spanning the last exons of M18BP1, was retrieved from the BAC clone using recombineering in EL350 bacteria. The mini-targeting plasmid was constructed by generating PCR fragments flanking the M18BP1 stop codon. These PCR fragments were cloned together with the floxed Neomycin selection cassette from pL452 (EcoRI-BamHI fragment) into pBluescript IISK+ using NotI, EcoRI, BamHI and SalI. In a subsequent cloning step, the EGFP tag was inserted with

EcoRI. For the final targeting vector, the 7 kb region was mini-targeted by recombineering with the NotI-Sall fragment containing the EGFP and floxed Neomycin selection cassette from the mini-targeting plasmid.

To generate M18BP1 knock-in cells, the NotI-linearized targeting vector was electroporated into feeder-independent wild type mES cells. Cells were selected in 180 µg/ml G418 (PAA) and 2 µM Ganciclovir (Invivogen). Single colonies were picked and screened by nested PCR to obtain the final mES cell clone (K1B2). Primers used for cloning and PCR screening are listed in Table S1.

Cell culture and transfections. BHK cells containing a *lac* operator repeat array³² were cultured in DMEM medium with 10% FCS and seeded on coverslips in 6-well plates for microscopy. After attachment cells were co-transfected with expression vectors for the indicated fluorescent fusion proteins and a LacI-GBP fusion^{33,34} using polyethylenimine (Sigma). After about 16 h cells were fixed with 3.7% formaldehyde in PBS for 10 min, washed with PBST (PBS with 0.02% Tween), stained with DAPI and mounted in Vectashield medium (Vector Laboratories).

Mouse ES cells were cultivated on gelatinized plates in High Glucose DMEM with L-Glutamine and sodium pyruvate, complemented with 15% FCS, β-mercaptoethanol, non essential amino acids (PAA), penicillin/streptomycin (PAA) and LIF in a 37°C incubator at 5% CO₂. For transfection with Lipofectamine 2000 (Invitrogen), mES cells were seeded on matrigel (BD Biosciences) coated coverslips.

HEK 293FT cells (Invitrogen) were cultivated on gelatinized plates in High Glucose DMEM with L-Glutamine and sodium pyruvate (PAA) complemented with 10% FCS, β-mercaptoethanol, non essential amino acids (PAA) and penicillin/streptomycin (PAA) in a 37°C incubator at 5% CO₂. The cells were transiently transfected one day after seeding using standard calcium phosphate transfection.

Microscopy. F3H samples were analyzed with a confocal fluorescence microscope (TCS SP5, Leica) equipped with a 63 × / 1.4 numerical aperture Plan-Apochromat oil immersion objective as described.³⁴ DAPI, EGFP and mCherry/RFP were excited by 405 nm diode laser, 488 nm argon laser and 561 nm diode-pumped solid-state laser, respectively. Images were recorded with a frame size of 512 × 512 pixels.

K1B2 cells were imaged using a Leica TCS SP5 confocal laser scanning microscope with a HCX PL APO CS 63x/1.3 NA glycerol immersion objective. Sequential excitation at 405 nm, 488 nm, 543 nm and 633 nm was provided by diode, argon and helium-neon gas lasers, respectively. Emission detection ranges of the photomultipliers were adjusted to avoid crosstalk between the channels. Maximum intensity projections of the confocal sections were generated using ImageJ software.

Intensity ratio measurement. Images were acquired with an IN Cell Analyzer 2000 (GE Healthcare) using a 40 × air objective and analyzed with the IN Cell Analyzer 1000 Workstation 3.7 (GE Healthcare). Green and red fluorescence intensities at the *lac* spots were quantified. After background subtraction, intensity ratios of red (prey) to green (bait) were calculated and plotted using Excel software (Microsoft).

Lentiviral knockdown and infection of K1B2 cells. Lentiviral shRNA sequences (Table S2) were selected from the TRC library³⁵ or designed using the TRC shRNA designer (<http://www.broadinstitute.org/rnai/public/>).

For lentiviral knock-downs one non-targeting shRNA and three shRNAs targeting mouse CENP-C mRNA (NCBI RefSeq NM_007683.3) were cloned into the lentiviral knock-down vector pLKOmod1³⁶ with MluI/XmaI.

For restricting lentiviral transduction to mouse cells, we replaced the commonly used VSVg protein during viral packaging with the ecotropic envelope protein of Moloney Murine Leukemia Virus. The new packaging vector pLP-ecoenv was generated by removing the VSVg sequence of pLP-VSVg (Invitrogen) by EcoRI digest, followed by T4 polymerase filling of the remaining vector and ligation of a EcoRI/NotI cut and T4 polymerase filled PCR product of the M-MLV ecotropic envelope sequence

(Primers: eco env fw 5'-CGAATTGCGCCGCCACCATGG CGCGTTCAACGCTCTCAAAA-3'; eco env rw 5'-TACGC GGCCGCTATGGCTCGTACTCTAT-3').

Lentiviral production was performed by seeding 4 million HEK293FT cells (Invitrogen) one day before transfection in gelatinized 10 cm dishes. On the following day, cells were transiently cotransfected with 8 µg psPAX2, 8 µg pLP-ecoenv and 8 µg of the respective pLKOmod1 vector using standard calcium phosphate transfection. Conditioned medium containing recombinant lentiviruses was harvested 48 h post transfection, aliquoted, snap frozen and stored at -80°C until further use.

K1B2 cells were transduced by seeding 3 × 10⁶ cells onto gelatinized 15 cm dishes containing mES cell medium supplemented with 4 µg/ml Polybrene (Sigma) and up to 20% conditioned virus medium. After 24 h the medium was replaced. 48 h post transduction, cells with stable integration of the pLKOmod1 vector were selected in mES cell medium containing 1.4–2 µg/ml puromycin (PAA) and then maintained in this selection medium until analysis.

Knock-down efficiency was determined at day five post infection by qRT-PCR and protein gel blotting. The following antibodies were used: CENP-C (Abcam ab50974), Suv4-20h2 (Hahn et al., in preparation).

RT-qPCR for monitoring M18BP1 and CENP-C expression levels. RNA of control and CENP-C knock-down cells was harvested at day 5 after transduction using RNeasy (Qiagen). 1.25 µg RNA was used for cDNA synthesis using Superscript III Kit (Invitrogen) and random hexameric primers (NEB). QPCR reactions were performed in technical triplicates using a Roche Light Cycler 480 with FAST SYBR[®] Master Mix (Applied Biosystems), and gene-specific primers (Table S3). Ct-values were normalized to the geometric mean of Actin and GAPDH for each individual cDNA and fold changes were calculated by the 2^{-ΔΔCt}-method.³⁷

Immunofluorescence. Immunofluorescence analyses were performed as described³⁸ using the following antibodies: CENP-A (C51A7, Cell Signaling Technology), H3S10P (06-570, Upstate) and Alexa 647 (A31573, Molecular Probes).

Plasmids. Encoding sequences of CENPs were amplified by PCR (Expand high fidelity^{PLUS} PCR System, Roche, Penzberg,

Germany). As forward primers we used 5'-GGGGACAAGT-TTGTACAAAAAGCAGGCTTCGAAAACCTGATTTTCAG-GGCGCCACC-3' as flanking regions followed by 20–26 bases of coding regions starting with 5'-ATGG-3' and as reverse primers we used 5'-GGGGACCACTTTGTACAAGAAAGCTGGGT-3' as flanking regions followed by 20–26 bases of coding sequences without stop codon. CENP encoding PCR fragments were transferred into vector pDONR221 by BP recombination reaction (Invitrogen, Carlsbad, CA, USA). After verification by DNA sequencing (MWG Biotech, Ebersberg, München, Germany), genes were transferred by LR recombination reactions to various modified pEGFP-C and pmCh-C (BD Biosciences, Clontech, Palo Alto, CA, USA) based destination vectors. The resulting expression vectors encode CENPs fused to the C-termini of EGFP and mCherry with SGTSLYKKAGFENLYFQGAT as linker sequence and TQLSCTKW added to the C-terminal ends of the FP-CENP fusions. Complete sequences are provided upon request. Correct full length expression of fusion constructs was confirmed by protein gel blots.

Full-length open reading frames of mouse M18BP1, the M18BP1 truncations and the mouse CENP-C truncations were PCR amplified from mouse cDNA derived from mES cells and cloned into the pDONR/Zeo GATEWAY entry vector (Invitrogen) using Gateway BP Clonase II enzyme mix (Invitrogen). PCR primers are listed in Table S4. Entry clones were recombined into target vectors pEGFP-N1-GW, pCMVmyc-GW and pGEX6P1-GW³⁹ using LR Clonase II enzyme mix (Invitrogen).

In vitro binding assays. Recombinant M18BP1 protein truncations (C1 aa1–440, C2 aa441–998, C3 aa325–800, C4 aa441–800, C5 aa735–800) were expressed as GST tagged versions in *E. coli* and purified on Glutathione-S-Sepharose (GE Healthcare). In vitro translation of CENP-C protein truncations (aa1–367, aa368–656, aa657–906) was performed using TnT[®] Quick Coupled Transcription/Translation System (Promega). 10 µl of the in vitro translated myc-tagged CENP-C and 5 µg M18BP1 GST-fusion protein coupled to Glutathione-S-Sepharose were incubated in IP buffer (50 mM Tris pH7.5, 150 mM NaCl, 1 mM EDTA, 0.1% NP40, 20% glycerol and proteinase inhibitor cocktail (Roche)) overnight at 4°C on a rotating wheel. The beads were washed four times with IP buffer containing 1 M NaCl and resuspended in 50 µl SDS loading

buffer (Roth). Bound proteins were separated on SDS polyacrylamidgels and detected by immunoblotting using α -myc antibody (9E10).

Co-immunoprecipitation in HEK293FT cells. HEK293FT cells (Invitrogen) were co-transfected with plasmids expressing EGFP-CENP-C and myc-M18BP1. Isolated nuclei were resuspended in high salt IP buffer (50 mM Tris pH 7.5, 500 mM NaCl, 1 mM EDTA, 0.1% NP40, 20% glycerol) with 4 strokes through a 19.5G syringe needle. After incubation on ice for 30 min the solution was sonicated 3x10³ at an amplitude of 30 in a Branson sonifier. The nuclear extract was diluted to a final concentration of 150 mM NaCl with no salt IP buffer and precipitates were removed by centrifugation. The extract was incubated overnight at 4°C on a rotating wheel with GFP-Trap beads (ChromoTek) and agarose beads. The beads were washed five times with IP buffer containing 300 mM NaCl and afterwards resuspended in SDS loading buffer (Roth). Proteins were separated on SDS-polyacrylamidgels and analyzed by protein gel blotting using α -myc (9E10) and α -GFP (Roche # 11814460001) antibodies.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We would like to thank S. F. Lichtenthaler and P.-H. Kuhn for generously providing the pLKMod1 vector. We thank N.G. Copeland for providing us with reagents for the bacterial recombineering system. We further thank A. Schmid and F. Büddefeld for help with cloning and C. Foster for critical reading of the manuscript.

G.S. and H.L. were supported by grants from DFG (TR-SFB) and BMBF (Episys). G.S. received additional support from Weigand'sche Stiftung. H.L. was further supported by the BioImaging Network and the Nanosystems Initiative Munich (NIM). SD wants to thank the DFG (SPP1128, SPP1395) and the TAB (2007 FE 9011) for support.

Supplemental Material

Supplemental material may be downloaded here:
<http://www.landesbioscience.com/journals/nucleus/article/18955/>

References

1. Foltz DR, Jansen LE, Black BE, Bailey AO, Yates JR, 3rd, Cleveland DW. The human CENP-A centromeric nucleosome-associated complex. *Nat Cell Biol* 2006; 8:458–69; PMID:16622419; <http://dx.doi.org/10.1038/ncb1397>
2. Obuse C, Yang H, Nozaki N, Goto S, Okazaki T, Yoda K. Proteomics analysis of the centromere complex from HeLa interphase cells: UV-damaged DNA binding protein 1 (DDB-1) is a component of the CEN-complex, while BMI-1 is transiently co-localized with the centromeric region in interphase. *Genes Cells* 2004; 9:105–20; PMID:15009096; <http://dx.doi.org/10.1111/j.1365-2443.2004.00705.x>
3. Okada M, Okawa K, Isobe T, Fukagawa T. CENP-H-containing complex facilitates centromere deposition of CENP-A in cooperation with FACT and CHD1. *Mol Biol Cell* 2009; 20:3986–95; PMID:19625449; <http://dx.doi.org/10.1091/mbc.E09-01-0065>
4. Guse A, Carroll CW, Moree B, Fuller CJ, Straight AF. In vitro centromere and kinetochore assembly on defined chromatin templates. *Nature* 2011; 477:354–8; PMID:21874020; <http://dx.doi.org/10.1038/nature10379>
5. Carroll CW, Milks KJ, Straight AF. Dual recognition of CENP-A nucleosomes is required for centromere assembly. *J Cell Biol* 2010; 189:1143–55; PMID:20566683; <http://dx.doi.org/10.1083/jcb.201001013>
6. Cohen RL, Espelin CW, De Wulf P, Sorger PK, Harrison SC, Simons KT. Structural and functional dissection of Mif2p, a conserved DNA-binding kinetochore protein. *Mol Biol Cell* 2008; 19:4480–91; PMID:18701705; <http://dx.doi.org/10.1091/mbc.E08-03-0297>
7. Sugimoto K, Kuriyama K, Shibata A, Himeno M. Characterization of internal DNA-binding and C-terminal dimerization domains of human centromere/kinetochore autoantigen CENP-C in vitro: role of DNA-binding and self-associating activities in kinetochore organization. *Chromosome Res* 1997; 5:132–41; PMID:9146917; <http://dx.doi.org/10.1023/A:1018422325569>

8. Screpanti E, De Antoni A, Alushin GM, Petrovic A, Melis T, Nogales E, et al. Direct binding of Cenp-C to the Mis12 complex joins the inner and outer kinetochore. *Curr Biol* 2011; 21:391-8; PMID:21353556; <http://dx.doi.org/10.1016/j.cub.2010.12.039>
9. Carroll CW, Silva MC, Godek KM, Jansen LE, Straight AF. Centromere assembly requires the direct recognition of CENP-A nucleosomes by CENP-N. *Nat Cell Biol* 2009; 11:896-902; PMID:19543270; <http://dx.doi.org/10.1038/ncb1899>
10. Bernad R, Sanchez P, Rivera T, Rodriguez-Corsino M, Boyarchuk E, Vassias I, et al. Xenopus HJURP and condensin II are required for CENP-A assembly. *J Cell Biol* 2011; 192:569-82; PMID:21321101; <http://dx.doi.org/10.1083/jcb.201005136>
11. Hemmerich P, Weidtkamp-Peters S, Hoischen C, Schmiedeborg L, Erliandri I, Diekmann S. Dynamics of inner kinetochore assembly and maintenance in living cells. *J Cell Biol* 2008; 180:1101-14; PMID:18347072; <http://dx.doi.org/10.1083/jcb.200710052>
12. Jansen LE, Black BE, Foltz DR, Cleveland DW. Propagation of centromeric chromatin requires exit from mitosis. *J Cell Biol* 2007; 176:795-805; PMID:17339380; <http://dx.doi.org/10.1083/jcb.200701066>
13. Schuh M, Lehner CF, Heidmann S. Incorporation of *Drosophila* CID/CENP-A and CENP-C into centromeres during early embryonic anaphase. *Curr Biol* 2007; 17:237-43; PMID:17222555; <http://dx.doi.org/10.1016/j.cub.2006.11.051>
14. Dunleavy EM, Roche D, Tagami H, Lacoste N, Ray-Galler D, Nakamura Y, et al. HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres. *Cell* 2009; 137:485-97; PMID:19410545; <http://dx.doi.org/10.1016/j.cell.2009.02.040>
15. Foltz DR, Jansen LE, Bailey AO, Yates JR, 3rd, Bassett EA, Wood S, et al. Centromere-specific assembly of CENP-a nucleosomes is mediated by HJURP. *Cell* 2009; 137:472-84; PMID:19410544; <http://dx.doi.org/10.1016/j.cell.2009.02.039>
16. Bergmann JH, Rodriguez MG, Martins NM, Kimura H, Kelly DA, Masumoto H, et al. Epigenetic engineering shows H3K4me2 is required for HJURP targeting and CENP-A assembly on a synthetic human kinetochore. *EMBO J* 2011; 30:328-40; PMID:21157429; <http://dx.doi.org/10.1038/emboj.2010.329>
17. Fujita Y, Hayashi T, Kiyomitsu T, Toyoda Y, Kokubu A, Obuse C, et al. Priming of centromere for CENP-A recruitment by human hMis18alpha, hMis18beta, and M18BP1. *Dev Cell* 2007; 12:17-30; PMID:17199038; <http://dx.doi.org/10.1016/j.devcel.2006.11.002>
18. Hayashi T, Fujita Y, Iwasaki O, Adachi Y, Takahashi K, Yanagida M. Mis16 and Mis18 are required for CENP-A loading and histone deacetylation at centromeres. *Cell* 2004; 118:715-29; PMID:15369671; <http://dx.doi.org/10.1016/j.cell.2004.09.002>
19. Maddox PS, Hyndman F, Monen J, Oegema K, Desai A. Functional genomics identifies a Myb domain-containing protein family required for assembly of CENP-A chromatin. *J Cell Biol* 2007; 176:757-63; PMID:17339379; <http://dx.doi.org/10.1083/jcb.200701065>
20. Pidoux AL, Choi ES, Abbott JK, Liu X, Kagansky A, Castillo AG, et al. Fission yeast Scm3: A CENP-A receptor required for integrity of subkinetochore chromatin. *Mol Cell* 2009; 33:299-311; PMID:19217404; <http://dx.doi.org/10.1016/j.molcel.2009.01.019>
21. Williams JS, Hayashi T, Yanagida M, Russell P. Fission yeast Scm3 mediates stable assembly of Cnp1/CENP-A into centromeric chromatin. *Mol Cell* 2009; 33:287-98; PMID:19217403; <http://dx.doi.org/10.1016/j.molcel.2009.01.017>
22. Leonhardt H, Rahn HP, Weinzierl P, Sporbert A, Cremer T, Zink D, et al. Dynamics of DNA replication factories in living cells. *J Cell Biol* 2000; 149:271-80; PMID:10769021; <http://dx.doi.org/10.1083/jcb.149.2.271>
23. Hsu JY, Sun ZW, Li X, Reuben M, Tatchell K, Bishop DK, et al. Mitotic phosphorylation of histone H3 is governed by Ipl1/aurora kinase and Glc7/PP1 phosphatase in budding yeast and nematodes. *Cell* 2000; 102:279-91; PMID:10975519; [http://dx.doi.org/10.1016/S0092-8674\(00\)00034-9](http://dx.doi.org/10.1016/S0092-8674(00)00034-9)
24. Lagana A, Dorn JF, De Rop V, Ladouceur AM, Maddox AS, Maddox PS. A small GTPase molecular switch regulates epigenetic centromere maintenance by stabilizing newly incorporated CENP-A. *Nat Cell Biol* 2010; 12:1186-93; PMID:21102442; <http://dx.doi.org/10.1038/ncb2129>
25. Boyer LA, Latek RR, Peterson CL. The SANT domain: a unique histone-tail-binding module? *Nat Rev Mol Cell Biol* 2004; 5:158-63; PMID:15040448; <http://dx.doi.org/10.1038/nrm1314>
26. Moree B, Meyer CB, Fuller CJ, Straight AF. CENP-C recruits M18BP1 to centromeres to promote CENP-A chromatin assembly. *J Cell Biol* 2011; 194:855-71; PMID:21911481; <http://dx.doi.org/10.1083/jcb.201106079>
27. Chung TL, Hsiao HH, Yeh YY, Shia HL, Chen YL, Liang PH, et al. In vitro modification of human centromere protein CENP-C fragments by small ubiquitin-like modifier (SUMO) protein: definitive identification of the modification sites by tandem mass spectrometry analysis of the isopeptides. *J Biol Chem* 2004; 279:39653-62; PMID:15272016; <http://dx.doi.org/10.1074/jbc.M405637200>
28. Mayya V, Lundgren DH, Hwang SI, Rezaul K, Wu L, Eng JK, et al. Quantitative phosphoproteomic analysis of T cell receptor signaling reveals system-wide modulation of protein-protein interactions. *Sci Signal* 2009; 2:ra46; PMID:19690332; <http://dx.doi.org/10.1126/scisignal.2000007>
29. Nousiainen M, Sillje HH, Sauer G, Nigg EA, Korner R. Phosphoproteome analysis of the human mitotic spindle. *Proc Natl Acad Sci USA* 2006; 103:5391-6; PMID:16565220; <http://dx.doi.org/10.1073/pnas.0507066103>
30. Olsen JV, Blagoev B, Gnani F, Macek B, Kumar C, Mortensen P, et al. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* 2006; 127:635-48; PMID:17081983; <http://dx.doi.org/10.1016/j.cell.2006.09.026>
31. Liu P, Jenkins NA, Copeland NG. A highly efficient recombineering-based method for generating conditional knockout mutations. *Genome Res* 2003; 13:476-84; PMID:12618378; <http://dx.doi.org/10.1101/gr.749203>
32. Tsukamoto T, Hashiguchi N, Janicki SM, Tumber T, Belmont AS, Spector DL. Visualization of gene activity in living cells. *Nat Cell Biol* 2000; 2:871-8; PMID:11146650; <http://dx.doi.org/10.1038/35046510>
33. Rothbauer U, Zolghadr K, Tillib S, Nowak D, Schermelleh L, Gahl A, et al. Targeting and tracing antigens in live cells with fluorescent nanobodies. *Nat Methods* 2006; 3:887-9; PMID:17060912; <http://dx.doi.org/10.1038/nmeth953>
34. Zolghadr K, Mortusewicz O, Rothbauer U, Kleinhaus R, Goehler H, Wanker EE, et al. A fluorescent two-hybrid assay for direct visualization of protein interactions in living cells. *Mol Cell Proteomics* 2008; 7:2279-87; PMID:18622019; <http://dx.doi.org/10.1074/mcp.M700548-MCP200>
35. Moffat J, Grueneberg DA, Yang X, Kim SY, Kloepper AM, Hinkle G, et al. A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell* 2006; 124:1283-98; PMID:16564017; <http://dx.doi.org/10.1016/j.cell.2006.01.040>
36. Kuhn PH, Wang H, Dislich B, Colombo A, Zeitschel U, Ellwart JW, et al. ADAM10 is the physiologically relevant, constitutive alpha-secretase of the amyloid precursor protein in primary neurons. *EMBO J* 2010; 29:3020-32; PMID:20676056; <http://dx.doi.org/10.1038/emboj.2010.167>
37. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, et al. 2002Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 2002; 3: H0034; PMID:12184808; <http://dx.doi.org/10.1186/gb-2002-3-7-research0034>
38. Lehnertz B, Ueda Y, Derijck AA, Braunschweig U, Perez-Burgos L, Kubicek S, et al. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr Biol* 2003; 13:1192-200; PMID:12867029; [http://dx.doi.org/10.1016/S0960-9822\(03\)00432-9](http://dx.doi.org/10.1016/S0960-9822(03)00432-9)
39. Schotta G, Sengupta R, Kubicek S, Malin S, Kauer M, Callen E, et al. A chromatin-wide transition to H4K20 monomethylation impairs genome integrity and programmed DNA rearrangements in the mouse. *Genes Dev* 2008; 22:2048-61; PMID:18676810; <http://dx.doi.org/10.1101/gad.476008>

Supplementary Material to:

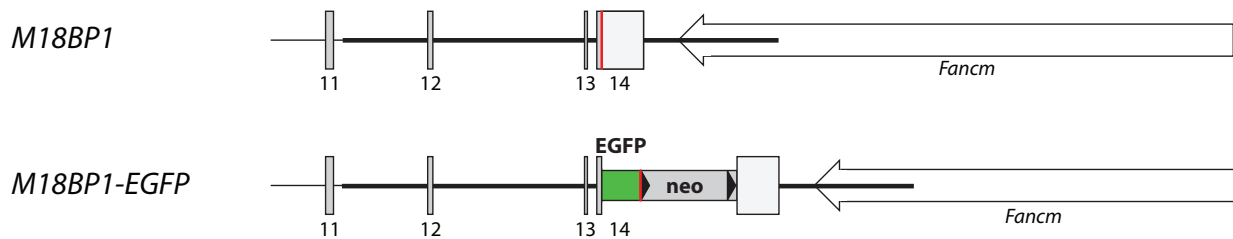
CENP-C facilitates the recruitment of M18BP1 to centromeric chromatin

Silvia Dambacher, Wen Deng, Matthias Hahn, Dennis Sadic, Jonathan Fröhlich, Alexander Nuber, Christian Hoischen, Stephan Diekmann, Heinrich Leonhardt and Gunnar Schotta
Nucleus 2012; 3(1): In Press. DOI: 10.4161/nucl.3.1.18955

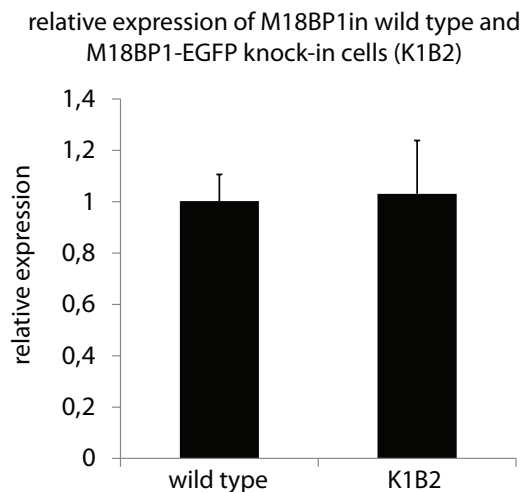
Supplementary Figure S1. Generation of M18BP1-EGFP knock-in cells.

(A) Targeting strategy for the M18BP1-EGFP knock-in allele. Part of the M18BP1 locus containing the last exons (11-14) is depicted in the scheme. The last exon contains the regular STOP codon (red line), followed by the 3' UTR. The targeting construct comprises the region shown by the thick black line. The knock-in allele contains the EGFP tag just before the regular STOP codon, followed by the loxP flanked Neomycin selection cassette. (B) RT-qPCR quantification of M18BP1 expression levels in the parental wild type mES cell line and the K1B2 cell line which carries the M18BP1-EGFP knock-in allele. Average expression levels from triplicate experiments, normalized to Actin and GAPDH are shown.

A

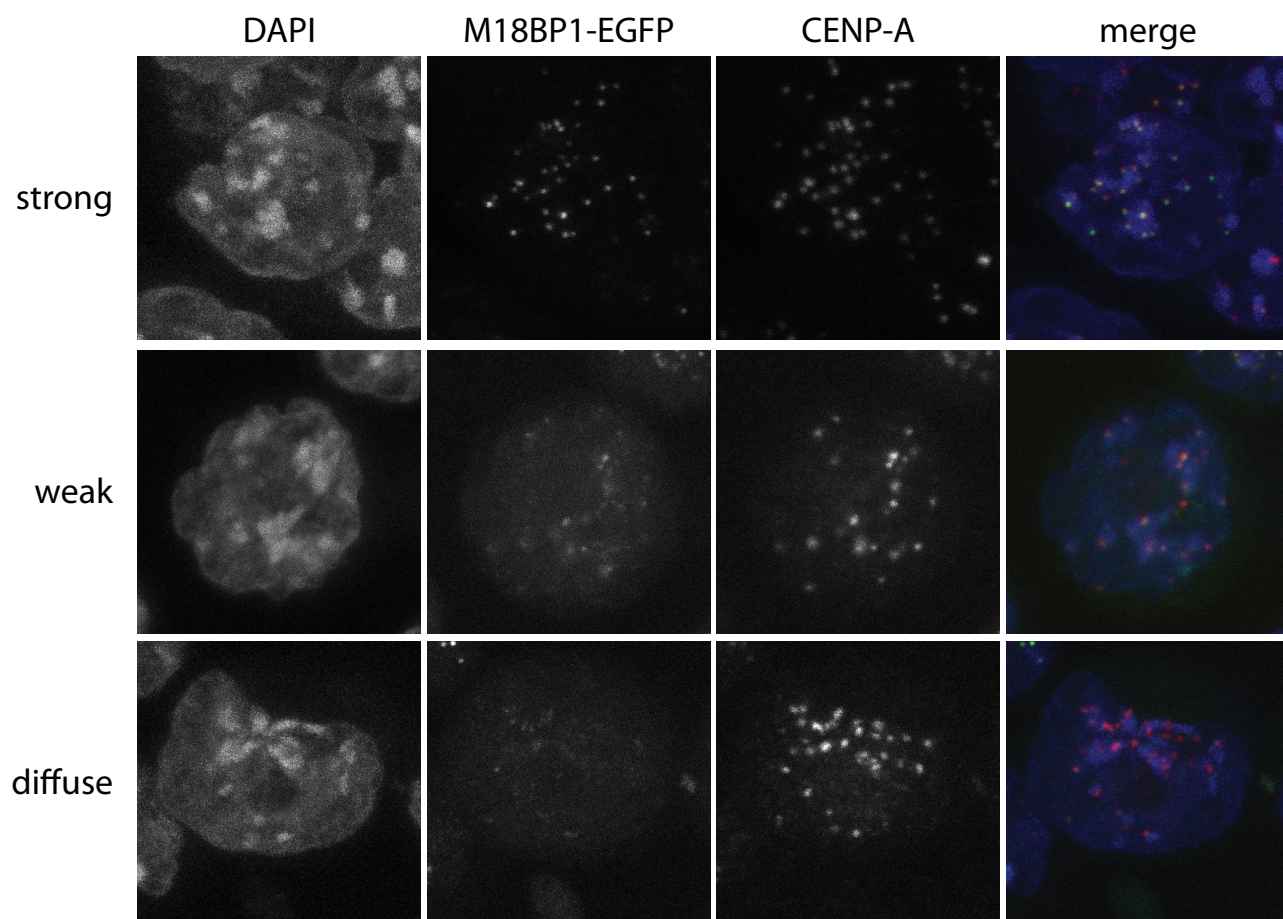


B



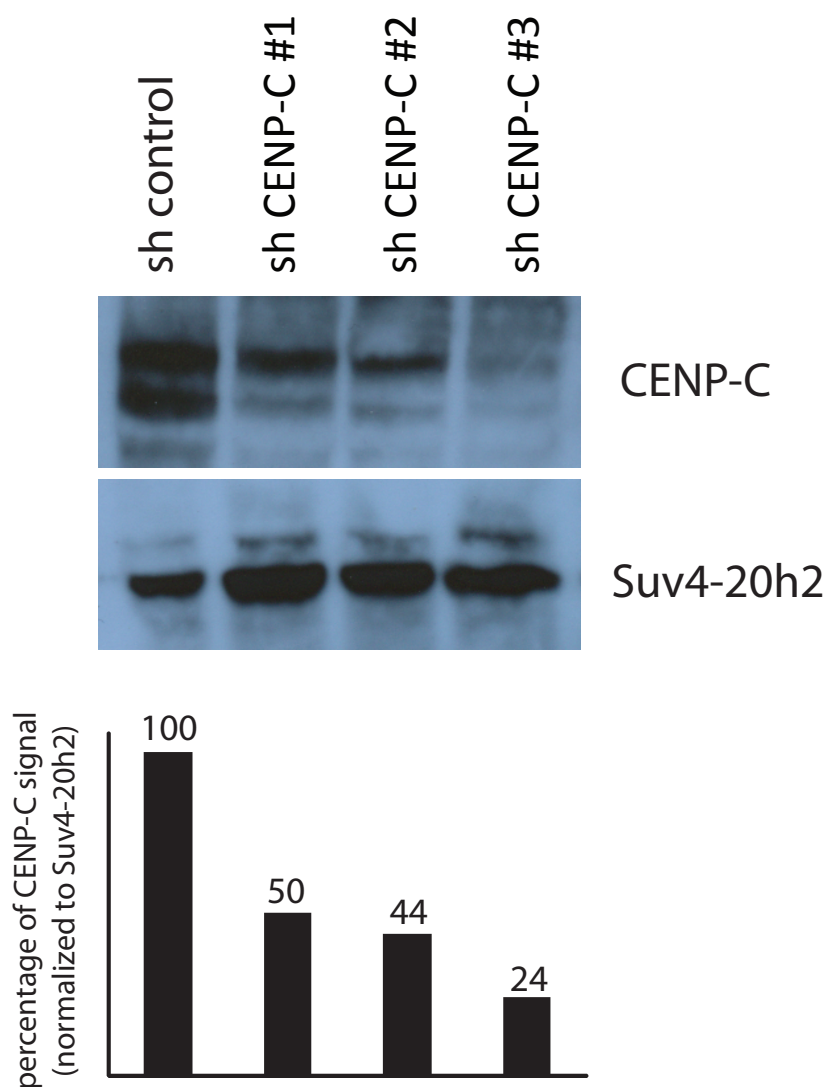
Supplementary Figure S2. Centromeric localization of M18BP1.

Localization of endogenously tagged M18BP1-EGFP in the K1B2 mES cell line. K1B2 cells were stained for CENP-A and confocal stacks were recorded. Maximum intensity projections of individual cells representing the different M18BP1 staining patterns: strong enrichment at centromeres (s), weak enrichment (w), no enrichment/diffuse nuclear (d) are shown.



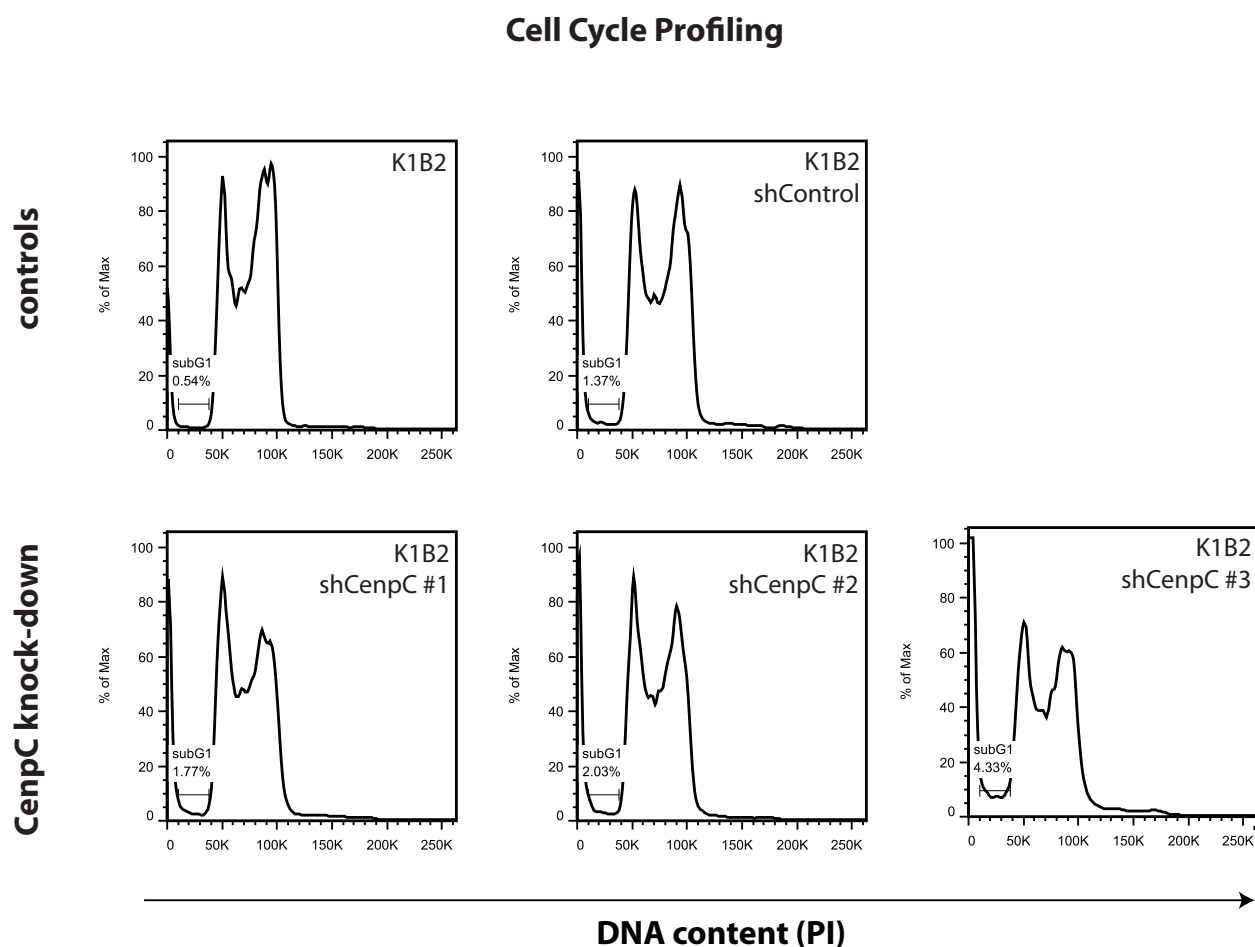
Supplementary Figure S3. CENP-C protein levels in control vs. CENP-C knock-down cells.

Five days after infection with the lentiviral knock-down vectors, control and CENP-C knock-down cells were harvested and nuclear extracts from these cells were probed for CENP-C, and Suv4-20h2 which served as loading control, by western blotting. Compared to shControl cells, all three knock-down oligos lead to reduced CENP-C protein levels, with shCENP-C #3 showing the strongest effect.



Supplementary Figure S4. Cell cycle profiles of K1B2 and CENP-C knock-down cells.

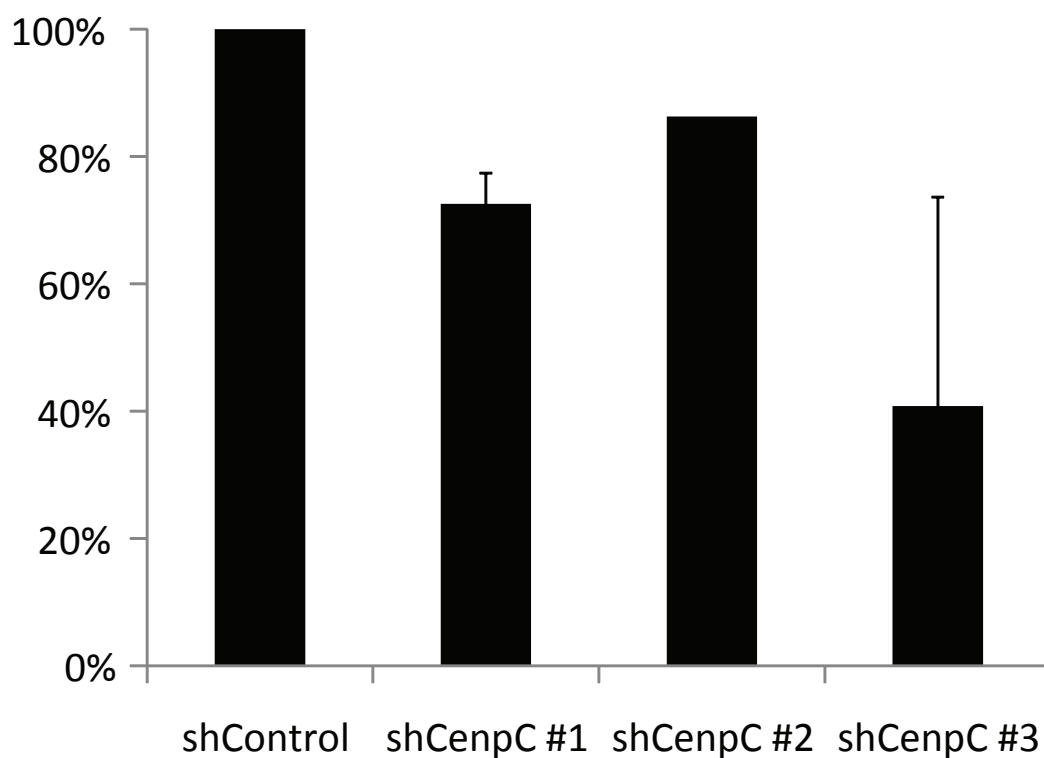
Five days after infection with the lentiviral knock-down vectors, control and CENP-C knock-down cells were harvested, fixed and stained for DNA content with propidium iodide (PI). FACS analysis revealed a normal cell cycle profile, however, the CENP-C knock-down cells tend to show higher levels of sub-G1 cells, indicative of cell death. This effect is most apparent with knock-down shCENP-C #3, which shows the strongest reduction of CENP-C mRNA levels.



Supplementary Figure S5. Relative cell numbers of shControl and shCENP-C knock-down populations.

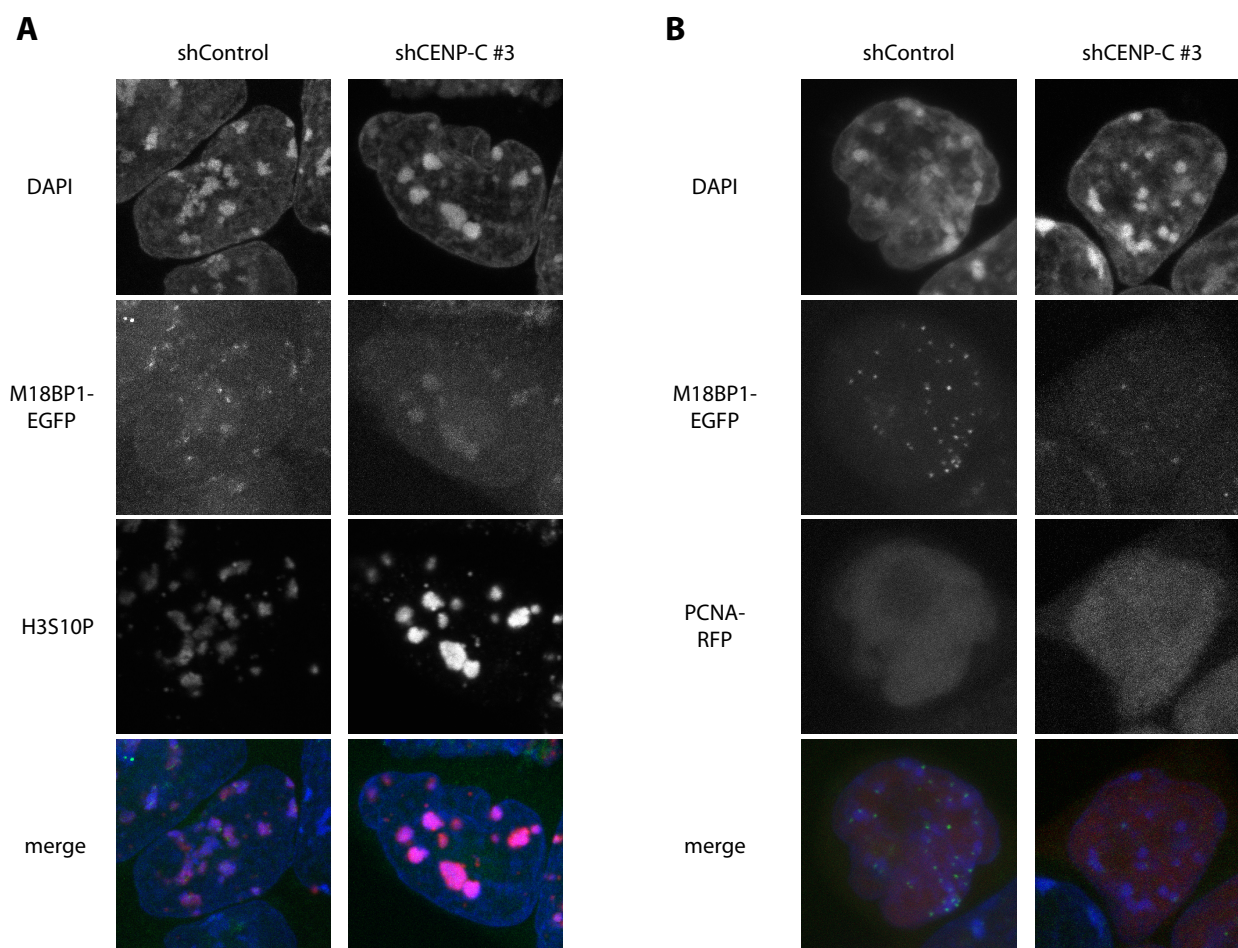
Five days after infection with the lentiviral knock-down vectors, control and CENP-C knock-down cells were harvested and counted. The cell number of the control population was set to 100%. The bar graph shows relative cell numbers for the individual shCENP-C knock-down populations.

**Relative number of surviving cells
5 days after transduction of CenpC
targeting shRNAs**



Supplementary Figure S6. Cell cycle analysis in CENP-C knock-down cells.

(A) Five days after infection with the lentiviral knock-down vectors, control and CENP-C knock-down cells were stained with H3S10P antibodies to detect cells in G2 phase. Maximum intensity projections of representative examples for G2 phase cells are shown. Control cells display a ‘weak’ centromeric staining pattern, in the CENP-C knock-down sample we observe cells which have lost the centromeric M18BP1 signals. (B) Four days after infection with the lentiviral knock-down vectors, control and CENP-C knock-down cells were transfected with PCNA-RFP expression plasmids. G1 cells which are characterized by small nuclei and diffuse PCNA signals were examined for M18BP1-EGFP localization. Control cells show ‘strong’ centromeric M18BP1-EGFP signals. In the CENP-C knock-down population we can detect G1 phase cells which have lost or strongly reduced centromeric enrichment of M18BP1.



Supplementary Table 1. Primers used for cloning the M18BP1-EGFP targeting construct.

primer	orientation	primer sequence 5'-3'
cloning fragment AB	f	ATAGCGGCCGCCTAACTCAAATGCAAAACC
cloning fragment AB	r	CGCAAGCTTTGATTAATAGTTTTTCACTAT
cloning fragment YZ	f	TCCAAGCTTCCATGACTTGCTCACCTTG
cloning fragment YZ	r	TGCACTAGTTATGAAAGAACTCTCATAATG
cloning fragment CD	f	ATAGCGGCCGCTCTCCACCACCAACACGGA
cloning fragment CD	r	TCGGAATTCGTCAGAATTGGAAAAGTAAT
cloning fragment EF	f	CGAGGATCCTGATAGACGACTTGCAGGAAT
cloning fragment EF	r	TCTGTGCGACAGTCCACAATCTTAACTCTG
nested PCR outer	f	ACCGCTTCCTCGTGCTTTAC
nested PCR outer	r	AAAGCCAAGCTCACTGTTTC
nested PCR inner	f	GATTGGGAAGACAATAGCAGGCATG
nested PCR inner	r	GCGCAAGTAAATCATCAAAAGGCTG

Supplementary Table 2. ShRNA oligonucleotide sequences.

name	species	direction	sequence – in 5' to 3' direction (targeted sequence in bold)	TRC code	targeted mRNA
non targeting shRNA	none	fw	CGCGTCCGGCAACAAGATGAAGAGCACCAACTCGAGT TGGTGCTCTTCATCTTGTGTTTTGGAAA	SHC202V	none
		rw	CCGGTTTCCAAAAACAACAAGATGAAGAGCACCAACT CGAGTTGGTGCTCTTCATCTTGTGTCGGGA		
shCENP-C1 #1	Mm	fw	CGCGTCCGGGCATGTTGGCCAAGATATATTCTCGAGA ATATATCTTGGCCAACATGCTTTTTGGAAA	from TRC designer	NM_0076 83.3
		rw	CCGGTTTCCAAAAAGCATGTTGGCCAAGATATATTCTC GAGAAATATATCTTGGCCAACATGCCCGGA		
shCENP-C #2	Mm	fw	CGCGTCCGGGTTTCGTCGATCTAATAGAATACTCGAGTA TTCTATTAGATCGACGAACTTTTGGAAA	from TRC designer	NM_0076 83.3
		rw	CCGGTTTCCAAAAAGTTCGTCGATCTAATAGAATACTC GAGTATTCTATTAGATCGACGAACCCGGA		
shCENP-C #3	Mm	fw	CGCGTCCGGGACATCACCGAATGTTTCATTTCTCGAGAA ATGAACATTCCGGTGATGTCTTTTTGGAAA	from TRC designer	NM_0076 83.3
		rw	CCGGTTTCCAAAAAGACATCACCGAATGTTTCATTTCTC GAGAAATGAACATTCCGGTGATGTCCCGGA		

Supplementary Table 3. Primers used for quantitative PCR.

gene	mRNA	direction	sequence in 5' to 3' direction
CENP-C	NM_007683.3	fw	aagccgacccatctcaatag
		rw	taagatccatggggacaagc
beta Actin	NM_007393.3	fw	ggcatcactattggcaacg
		rw	tccataccaagaaggaagg
M18BP1	NM_172578.2	fw	ctccaaaaggccagcatcacg
		rw	ttgccggaggtaggctgtcc
GAPDH	NM_008084.2	fw	tcaagaaggtggtgaagcag
		rw	gttgaagtcgcaggagacaa

Supplementary Table 4. Primers used for cloning of M18BP1 and CENP-C truncation constructs.

gene	mRNA	direction	sequence in 5' to 3' direction
M18BP1	NM_172578.2	fw	ggggacaagttgtacaaaaagcaggcttaactatgattgtaacaccttga
		rw	ggggaccactttgtacaagaaagctgggtcgtcagaattggaaaagtaa
M18BP1 (aa1-440)	NM_172578.2	fw	ggggacaagttgtacaaaaagcaggcttaactatgattgtaacaccttga
		rw	ggggaccactttgtacaagaaagctgggtctcttctgtttctgtctg
M18BP1 (aa441-998)	NM_172578.2	fw	ggggacaagttgtacaaaaagcaggcttaactatgcaggaaacagcaagag
		rw	ggggaccactttgtacaagaaagctgggtcgtcagaattggaaaagtaa
M18BP1 (aa325-800)	NM_172578.2	fw	ggggacaagttgtacaaaaagcaggcttaactatgactgttgtaaagaag
		rw	ggggaccactttgtacaagaaagctgggtcatgtttcgggacacctgg
M18BP1 (aa441-800)	NM_172578.2	fw	ggggacaagttgtacaaaaagcaggcttaactatgcaggaaacagcaagag
		rw	ggggaccactttgtacaagaaagctgggtcatgtttcgggacacctgg
M18BP1 (aa735-800)	NM_172578.2	fw	ggggacaagttgtacaaaaagcaggcttaactatggaccatctacctggtt
		rw	ggggaccactttgtacaagaaagctgggtcatgtttcgggacacctgg
CENP-C (aa1-367)	NM_007683.3	fw	ggggacaagttgtacaaaaagcaggcttaactatggcctCgttcacatctggatc
		rw	ggggaccactttgtacaagaaagctgggtctttattttcaggagatcgacaa
CENP-C (aa368-656)	NM_007683.3	fw	ggggacaagttgtacaaaaagcaggcttaactatgcaatctgagactgccaaaac
		rw	ggggaccactttgtacaagaaagctgggtcttcataattcttgaacctggaag
CENP-C (aa657-906)	NM_007683.3	fw	ggggacaagttgtacaaaaagcaggcttaactatgccaggagcagtaattctg
		rw	ggggaccactttgtacaagaaagctgggtccctttttatttgagtaaaaagaag

3. Discussion

Heterochromatin has important functions for gene regulation and maintenance of genomic integrity, thus heterochromatin dysregulation can result in severe diseases (Hahn et al. 2010). The underlying epigenetic mechanisms of heterochromatin formation and regulation have been studied intensively in the last decades, but still many open questions remain to be answered. The goal of this dissertation was to identify and describe new mechanisms that regulate centromeric and pericentromeric heterochromatin during organismal development. The results were presented in two review articles and four papers, all in peer-reviewed scientific journals.

3.1 Suv4-20h enzymes and H4K20me3 in embryonic development

Heterochromatin formation and its relevance for developmental gene regulation has been a subject of detailed studies in the recent years. However, these analyses mostly emphasized on TrxG- and PcG-mediated control of developmental transitions. One focus of this PhD thesis was to identify which roles Suv4-20h enzymes and the pericentromeric heterochromatin mark H4K20me3 might play in the regulation of developmental processes. In this context, *Xenopus laevis* was a perfect model organism offering an efficient tool kit for manipulating and studying embryogenesis. Suv4-20h enzymes and H4K20 methylation are well known for contributing to the establishment of pericentric heterochromatin (Lachner et al. 2004; Schotta et al. 2004; Schotta et al. 2008). Furthermore they have been implicated in contributing to cellular processes like general transcriptional regulation (Kapoor-Vazirani et al. 2011; Magklara et al. 2011). However, in *Nicetto et al. 2013*, we could show a novel function of Suv4-20h-dependent H4K20me3, namely the specific regulation of the pluripotency gene XOct-25. This is the *Xenopus* homolog of the mammalian Oct4 gene, a master regulator of stem cell self-renewal and differentiation. Morpholino knock-down experiments of the X(enopus)Suv4-20h1 and XSuv4-20h2 homologs resulted in defective eye and melanocyte development. Those phenotypes are indicative of failures in neuroectodermal differentiation. By closer examination of global transcription changes in XSuv4-20h-depleted embryos the pluripotency-related XOct-25, was identified as one of the top10 upregulated genes. Our analyses showed that XSuv4-20h enzymes are required for the restriction of XOct-25 expression to the sensorial cell layer of the ectoderm during gastrulation. A combined depletion of XOct-25 and XSuv4-20h could however rescue the morphological defects and to a large extent the misregulated neuronal gene expression. These results imply, that XSuv4-20h and X-Oct25 may act in the same pathway. ChIP

experiments verified that the 5'UTR region of XOct-25 was enriched with H4K20me3, whereas in XSuv4-20h morphants, H4K20me3 levels were significantly decreased. These findings suggest that XOct-25 is a direct target of XSuv4-20h regulation. To address the question, whether the genetic interaction between Suv4-20h and XOct-25/Oct4 exists in mammals as well, we analyzed Oct4 levels and the differentiation behaviour of murine ESCs lacking Suv4-20h enzymes. It was described before, that mammalian Oct4 is not only silenced by transcriptional control mechanisms but also by epigenetic silencing mechanisms like DNA methylation (Li et al. 2007), histone deacetylation (Feldman et al. 2006) or G9a-mediated H3K9me2 (Tachibana et al. 2002; Feldman et al. 2006). However, no link to Suv4-20h-mediated H4K20me3 regulation was known. We could demonstrate that mouse ES cells with a Suv4-20h1 and Suv4-20h2 double knock-out have higher Oct4 levels as wild type cells. We used embryoid body (EB) formation as a system to check Oct4 expression levels in differentiation and to test whether abrogation of Suv4-20h results in defective lineage specification. Embryoid bodies are spontaneously formed colonies of ES cells upon withdrawal of LIF (Keller 1995). They can undergo differentiation into all three germ layers and are therefore commonly used to test pluripotency and differentiation behaviour of ESC lines. We could demonstrate that after differentiation, Suv4-20h-mutant EBs were smaller in size and showed higher Oct4 levels as control cells. Furthermore, unlike the wild type EBs, Suv4-20h DKO EBs didn't form autonomously beating regions, indicating defects in cardiomyocyte formation. In addition, we established a qPCR marker analysis, which showed enhanced induction of FoxA2, and downregulation of Gata4 in Suv4-20h DKO ES cells. This shift in mesendoderm marker expression and the perturbed formation of functional cardiomyocyte demonstrated the comprised differentiation capacity of Suv4-20h DKO ESCs. Altogether our results emphasize the importance of Suv4-20h-dependent XOct-25 and mouse Oct4 regulation in accurate germ layer specification and cell differentiation. These findings are similar to previously published data showing that Oct4 suppresses neuroectodermal and promotes mesendodermal differentiation in mouse ESCs (Thomson et al. 2011). We therefore hypothesize that Suv4-20h enzymes are contributing, at least partially, to germ layer-specification of ESCs.

3.2 H3K56me3: a novel pericentric heterochromatin mark

As we have shown in chapter 2.3, histone modifications can play essential roles in differentiation. They are also accounting for various other processes such as DNA damage repair or replication. Deciphering the complex code of covalent histone

modifications is challenging, since the number and types of modifications is huge and still new ones are being discovered (Tan et al. 2011). It is critical to learn, how those modifications are generated, how they can be controlled and what impact they have on the living cell. One recently discovered histone modification is methylation of H3K56, a histone core residue (Garcia et al. 2007). H3K56 is a residue which is known to play roles in DNA damage response (Masumoto et al. 2005; Vempati et al. 2010; Wurtele et al. 2012) and chromatin integrity (Celic et al. 2006; Driscoll et al. 2007). H3K56me1 is generated by the G9a methyltransferase (Yu et al. 2012) which also accounts for H3K9me1 and H3K9me2 (Tachibana et al. 2002) possibly due to the conserved lysine-serine-threonine (K-S-T) motif in histone H3. However the mechanisms controlling H3K56me3 and its functional relevance remained unclear. We have addressed this question, by generation of a H3K56me3 antibody, verifying its specificity and investigating its localization in different microscopy experiments. Our data show, that H3K56me3 is an evolutionarily conserved modification. It is enriched at pericentromeric heterochromatin throughout the cell cycle, except the S phase. The reduced H3K56me3 levels in this cell cycle stage might be explained by a masking of the trimethylation by binding of protein interactors. However, we were not successful in isolating consistent binders of H3K56me3 in peptide pull down experiments. An active removal of H3K56 trimethylation during replication might also account for the reduced levels in S phase. Furthermore, H3K56 is a target of acetyltransferases, that modify this residue in S phase, before histones are integrated into DNA (Recht et al. 2006; Rufiange et al. 2007; Williams et al. 2008). Newly synthesized histones carrying H3K56ac could therefore simply exchange H3K56me3 histones in the newly replicated DNA. The K-S-T motif at the H3K56 position is target of the H3K9 histonemethyltransferase G9a. For this reason we assumed that H3K9me3-specific KMTs and KDMs could also target H3K56. Consistent with that idea, we found a loss of H3K56me3 in Suv39h knock-out mouse fibroblasts, accompanied by an increase in H3K56me1. These findings demonstrated that it is highly likely that Suv39h enzymes mediate trimethylation of H3K56. H3K56me1 was shown to bind to PCNA and to be crucial for DNA replication in S phase (Yu et al. 2012). This need of H3K56me1 for faithful replication would support the hypothesis, that it is important to control the H3K56me3 state, that does not allow PCNA binding. To test, if the H3K56me3 mark can be actively removed by a member of the jmjC-class of histone demethylases [reviewed in (Pedersen and Helin 2010)], we generated a GFP-tagged expression vector collection from a pDONR library of jmjC proteins (Fodor et al. 2006). The over-expression of the Jmjd2 family of

proteins in mouse and human cells resulted not only in a loss of H3K9me3 (Fodor et al. 2006) but also H3K56me3, likely due to the conserved K-S-T motif. A notable observation was, that the yet uncharacterized hKDM4DL (homolog to mouse pseudogene JmjD2e) is strongly expressed in testis and can demethylate H3K9me3 and H3K56me3 residues. This demonstrates clearly, that histone modifying enzymes may have more than one target, a finding which is supported by a recent paper showing another heterochromatin mark H3K64me3 being Suv39h-dependent (Daujat et al. 2009). H3K56me3 is localizing to pericentromeric heterochromatin in interphase nuclei as H3K9me3 and H4K20me3 also do. However in mitotic chromosome spreads H3K56me3 and H4K20me3 localizations differ, which implies disparate functions and will be interesting to analyze in subsequent studies. Our data furthermore show, that H3K56me3 exists in *C.elegans*, where it co-localizes with H3K9me3-positive regions and seems to be regulated by some but not all H3K9-specific KMTs. Although we could not describe the functional relevance of H3K56me3, our findings open new possibilities for further studies. Standard approaches like knock-downs of the responsible KMTs and KDMs however, will be hard to evaluate, since these would also affect the H3K9 and H3K64 methylation states, making it hard to allocate effects to one single histone modification.

3.3 Suv4-20h enzymes in heterochromatin organization

Proper heterochromatin formation is important for stabilizing genomic integrity and gene expression programs and upon dysregulation can result in severe diseases. However, heterochromatin is not simply a set of chromatin-associated proteins, but rather a site of dynamic interactions of a network of heterochromatin proteins. We could show that Suv4-20h2 is an important part of this network regulating chromatin compaction and cohesin recruitment. We established a fluorescence recovery after photobleaching (FRAP) protocol that enabled us to analyze GFP-tagged heterochromatin protein mobilities in MEFs and mouse ESCs. The proteins were transiently expressed and cells with similar fluorescence intensities were selected for FRAP experiments. This was necessary, since the protein-saturation of possible binding sites could falsify our measurements. We found HP1 α being a relatively mobile protein, a finding in agreement with previous studies (Cheutin et al. 2003). Suv39h1 and Suv4-20h1 proteins showed intermediate recovery kinetics whereas Suv39h2 and Suv4-20h2 associate extremely stable with heterochromatin. Notably, we could also demonstrate heterochromatin proteins being less mobile in MEF cells, as compared to ESCs, which is consistent with earlier studies (Meshorer et al. 2006).

This hyperdynamic binding of proteins to chromatin is a hallmark of pluripotent ESCs and is suggested to contribute to chromatin plasticity in undifferentiated ESCs. To exclude overexpression-artifacts in the mobility determinations, fluorescence correlation spectroscopy analyses (FCS) (Erdel et al. 2010) with a Suv4-20h2-EGFP knock-in ESC line was performed. This knock-in cell line expresses Suv4-20h2 from the endogenous genomic locus thus minimizing overexpression-artifacts. The FCS diffusion coefficient measurements showed that Suv4-20h2 binds more tightly to chromatin as HP1 α and HP1 β . In addition, the immobile pool of Suv4-20h2 was about 10 times higher than that of HP1. These FCS experiments could confirm the plausibility of our FRAP data. A recent study of Suv4-20h2 demonstrated its stable heterochromatin association (Souza et al. 2009), however the detailed binding mechanism remained an open question. In our manuscript, we could clearly show that a small region in the C-terminal part of Suv4-20h2 harbors multiple interaction sites for all three HP1 isoforms. Since HP1 levels at heterochromatin are enriched the binding to this short region in the C-terminus of Suv4-20h2 can explain the stable heterochromatin association. We therefore termed it „clamp“ domain. By measuring mobilities of truncated Suv4-20h2-GFP proteins in wild type cells and MEFs lacking HP1 α , we confirmed the relevance of HP1 for stable heterochromatin localization. We hypothesized that Suv4-20h2 might be a structural component of heterochromatin, bridging H3K9me3/HP1-rich chromatin regions. Thus we tested cells lacking Suv4-20h for alterations in chromatin organization. Our biochemical assays and 3D-SIM super resolution microscopy data could demonstrate that Suv4-20h2 is a central part of the heterochromatin network regulating nuclear architecture. Chromatin in Suv4-20h knock-out cells was more accessible to micrococcal nuclease (MNase). Interestingly, the clamp-domain of Suv4-20h2, was sufficient to restore chromatin compaction in the MNase accessibility experiments, suggesting that the enzymatic activity of Suv4-20h2 is dispensable for chromatin compaction. Two very interesting findings of these experiments may open up new possibilities in examining the roles of Suv4-20h2 in chromatin architecture. Firstly, the MNase digests in Suv4-20h mutant cells showed, in addition to the opening of chromatin, also alterations in nucleosomal repeat length. This is an indication, that abrogation of Suv4-20h2 may affect linker histone H1 association with chromatin, since a similar phenotype was detected in cells that lack linker histone H1 (Fan et al. 2005). Secondly, 3D-SIM analysis showed strong chromatin compaction around nucleoli, chromocenters and the nuclear periphery upon strong Suv4-20h2 over-expression. Moreover our data in Suv4-20h mutant cells revealed a perturbation in nuclear pore arrangement. One of

the top hits in our mass-spec analysis of Suv4-20h2 interactors was LaminB1. Nuclear pore proteins have been described to interact with A- and B-type lamins (Al-Haboubi et al. 2011) and a recent publication demonstrated the binding of the lamin-B receptor to H4K20me2 (Hirano et al. 2012). It will be interesting to follow this lead to further elucidate the role of Suv4-20h proteins in organizing nuclear envelope to chromatin interactions. Following the idea of Suv4-20h2 being a crucial component in ensuring the integrity of nuclear structure, our finding, that Suv4-20h-deficient cells display increased number of chromocenters is of notable interest. Similar observations have been described recently in Prdm3 and Prdm16 knockdown cells. A double knock-down of both Prdms resulted in an extreme chromocenter scattering phenotype (Pinheiro et al. 2012), indicating that the loss of Prdm3 and Prdm16 interferes in a rather initial and elementary step of chromocenter establishment. The chromocenter scattering phenotype in Suv4-20h mutant fibroblasts is Suv4-20h2-dependent, as it can be rescued by re-expression of the full length protein. However, the expression of the clamp domain could not restore the phenotype. Thus we assume that additional proteins, that might bind outside of the clamp domain, are necessary for chromocenter clustering. Accordingly the chromocenter scattering in Suv4-20h mutant cells could be a consequence of reduced histone H1 levels, as histone H1 variants participate in higher order heterochromatin formation (Lu et al. 2009b). In consequence, future studies should elaborate on the relation of Suv4-20h2 and histone H1 and their functional relevance for heterochromatin assembly. Defective heterochromatin can cause genomic instability (Peters et al. 2001). In fact, we could detect an increase in mitotic defects in Suv4-20h mutant fibroblasts, which could be rescued by Suv4-20h2 re-expression. An interesting observation was the increased centromere distance in mitotic Suv4-20h mutant chromosomes, which could be rescued by full length Suv4-20h2. To gain more insight into a possible role of Suv4-20h2 in sister chromatid cohesion, we wanted to identify possible Suv4-20h2 interacting proteins. Mass-spectrometric analyses of Suv4-20h2 binders and Suv4-20h2-IP experiments identified cohesin subunits Smc1 and Smc3 as interacting proteins. These findings indicated a potential link of Suv4-20h2 and the cohesin complex. Cohesin is connecting sister chromatids at pericentromeric regions until they are separated in anaphase. Cohesin loading at euchromatin seems to be mediated by CTCF chromatin insulator or mediator complex proteins (Parelho et al. 2008; Kagey et al. 2010). However the loading mechanisms of cohesin to pericentromeric heterochromatin were yet unclear. ChIP data demonstrated that cells lacking Suv39h or Suv4-20h enzymes show dramatically reduced Rad21/Smc3

levels at major satellite repeats. To exclude clonal effects, we had to analyse biological replicates of all ChIP experiments. MEF cell lines, that were isolated and established from different embryos were used to confirm the data. Thus our results imply a substantial function of the Suv39h-Suv4-20h pathway in cohesin recruitment to pericentromeric heterochromatin.

The finding that Suv39h mutant cells display a loss of cohesin at Major Satellite Repeats may appear contradictory to a previous study at first glance. (Koch et al. 2008) could not demonstrate loss of cohesin in Suv39h mutant MEF, however this could be explained by the semi-quantitative ChIP assay they used, which is not suitable for highly repetitive sequences like mouse major satellites. Strikingly, we could rescue pericentric cohesin localization in Suv4-20h deficient cells, not only with the full length Suv4-20h2 protein but also with the clamp domain alone. Cell cycle synchronized ChIP of cohesin subunits should elucidate, whether Suv4-20h2 is important for initial cohesin recruitment or for its maintenance at heterochromatin. Wild type and Suv4-20h DKO cell lines have been arrested in G0, S-phase or mitosis. In G0 cells we detected a complete loss of cohesin from pericentromeric heterochromatin, in the later cell cycle stages cohesin levels increased indicating Suv4-20h independent recruitment pathways might exist. Again the observed loss of cohesin could be rescued by the clamp domain or by full length Suv4-20h2 confirming its necessity for proper cohesin loading. Since the clamp domain can interact with cohesin and compact chromatin we cannot affirm whether chromatin compaction is necessary for cohesin recruitment or not. Based on recent publications and our data, we propose the following sequential pathway of cohesin recruitment to pericentric heterochromatin (Figure 5).

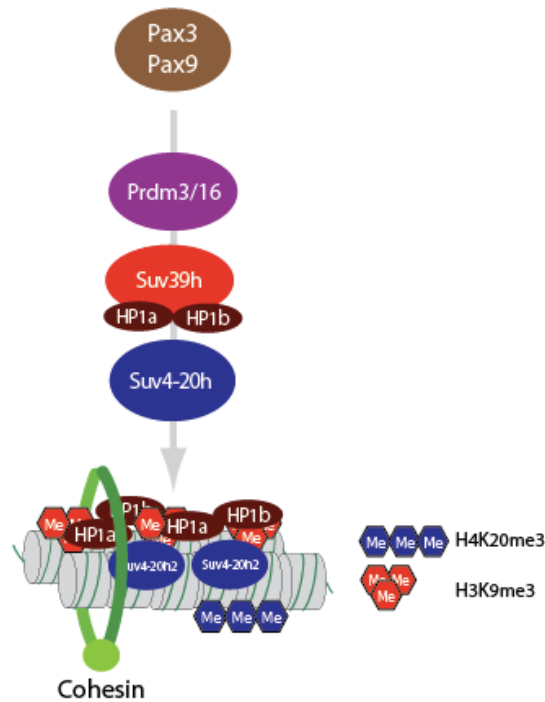


Figure 5. Pericentric cohesin recruitment

Mouse pericentric heterochromatin consists of major satellite repeats. The transcription factors Pax3 and Pax9 have been defined as regulators of mouse pericentromeric heterochromatin. They can regulate the RNA output from the major satellite repeats (Bulut-Karslioglu et al. 2012) which is important for proper heterochromatin establishment. Prdm3 and Prdm16 have been demonstrated to play essential roles in establishing H3K9me1 (Pinheiro et al. 2012) which is critical for subsequent H3K9me3 via the Suv39h KMTs. HP1 proteins bind to H3K9me3 and recruit Suv4-20h enzymes, that generate H4K20me3. Suv4-20h2 localization at heterochromatin is necessary for correct cohesin recruitment and chromatin architecture.

Our cell cycle ChIP analysis of cohesin subunits demonstrated that Suv4-20h2-dependent cohesin loading occurs in G0. While the cell progresses through the cell cycle, other mechanisms gradually recruit cohesin. In Suv4-20h mutants however, a full cohesin loading to pericentromeric chromatin, comparable to WT cells is not possible. Taken into account, that already a small reduction of heterochromatin-associated cohesin results in anomalies of mitotic fidelity (Eckert et al. 2007) the mild mitotic defects that we observed in Suv4-20h knock-out cells are comprehensible. Dysregulation of heterochromatin can result in genomic instability (see Figure 6), which is, together with cohesin defects (Thompson et al. 2010) a hallmark of cancer. Reduced H4K20me3 levels, an indirect measure of Suv4-20h2 presence, have been correlated with a variety of human cancers (Fraga et al. 2005; Tryndyak et al. 2006; Van Den Broeck et al. 2008). We therefore assume that Suv4-20h-dependent cohesin recruitment is crucial to guarantee genomic integrity, and if dysregulated, contributes to the transition of a „normal“ cell to cancer cell. The negative survival

prognosis of tumors with low H4K20me3 levels supports this hypothesis (Schneider et al. 2011).

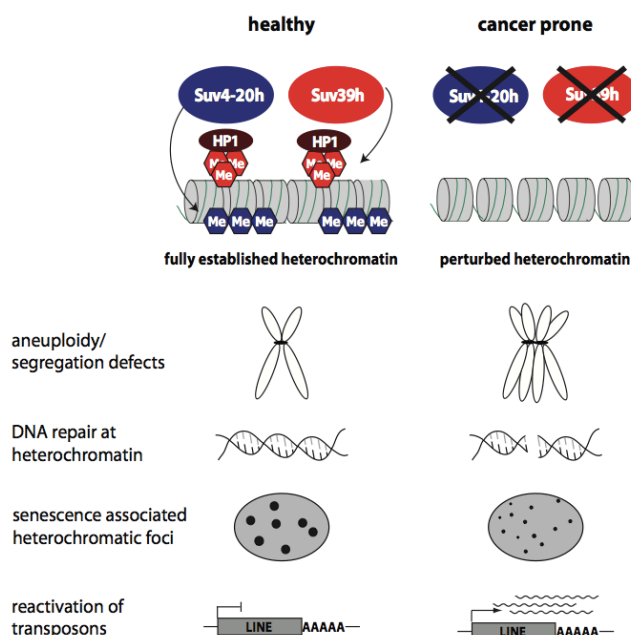


Figure 6. Heterochromatin dysregulation can increase the risk for developing cancer

Heterochromatin has important functions in gene regulation and maintenance of genomic stability. Dysregulation of heterochromatin-associated modifications, or loss of structural components like Suv4-20h2 and cohesin, might perturb these functions, resulting in chromosome segregation defects, inefficient DNA damage repair, failure to enter the senescence program and reactivation of mobile elements. All these defects increase the risk for a “normal” cell to transform into a cancer cell.

3.4 Heterochromatin in centromere formation

Chromosomes have to be accurately distributed in mitosis, to assure genomic integrity. Assembly of the kinetochore, that directs mitotic spindle attachment, happens at the centromeres. The processes in centromere establishment are poorly understood in vertebrates. To which extent the underlying DNA sequence is really important for centromere establishment (Harrington et al. 1997) or not (Voullaire et al. 1993; Kapoor et al. 1998) is under discussion. However epigenetic factors are nowadays more and more believed to play major roles in centromere propagation. The incorporation of the histone H3 variant Cenp-A is a critical step in the formation of functional centromeres (Regnier et al. 2005). Centromeric chromatin is marked by Cenp-A nucleosomes, together with nucleosomes containing H3K4me2, thereby differing from the neighboring pericentromeric heterochromatin that carries mainly H3K9me3 and H4K20me3 marks. M18bp1/Knl2 a member of the Mis18 complex, was demonstrated to be crucial for centromere formation and Cenp-A loading

(Maddox et al. 2007). In *Dambacher et al. 2012*, we generated and analyzed endogenous M18bp1-EGFP-tag knock-in mouse ESCs, to learn more about its biological functions. With this cell line it was possible for the first time, to study cell cycle-dependent localization and protein-protein interactions of the endogenous protein. Hence, we could exclude overexpression artifacts that other expression systems would induce. We could demonstrate a cell cycle specific association of endogenous M18bp1 with chromatin. For these experiments, we had to test several ways of attaching and growing living ESCs onto microscope cover slips. Coating of cover slips with Matrigel turned out to be the most efficient way to cultivate the cells. Using this protocol, we could describe M18bp1 localization and enrichment at centromeric chromatin from telophase to G1 cell cycle stage. This is consistent with data from earlier studies (Fujita et al. 2007), where the human M18bp1 protein showed a similar localization. In addition, we could describe an *in vivo* interaction of M18bp1 with the centromere protein Cenp-C. Our data demonstrated, that the middle part of M18bp1 is important for the interaction with Cenp-C. This region of the protein contains a SANT domain, a module of ~50 amino acids that can also be found in many subunits of chromatin remodelling complexes (Boyer et al. 2004). The SANT domain can bind histones, however not much is known about its functions. Testing the different M18bp1 truncations for interaction with Cenp-C in the F3H assay revealed that the SANT domain alone could not recruit Cenp-C. This might be due to an interference of the C-terminal EGFP tag with proper protein folding or that additional regions of the M18bp1 middle part are necessary for the interaction. With *in vitro* recombinant protein binding experiments we could show that the C-terminal part of Cenp-C mediates the M18bp1 interaction. This finding is in agreement with recently published work, demonstrating that *X.laevis* and human M18bp1 homologs can directly bind to human Cenp-C *in vitro* (Moree et al. 2011). To analyze the role of Cenp-C in a functional context, we used a lentiviral knock-down system that was established in our lab. After Cenp-C knockdown in the M18bp1-EGFP knock-in ESCs, we found a reduction of centromeric M18bp1 localization, independent of the cell cycle stage. Interestingly, the microscopic analysis of transient overexpression of Cenp-C in the M18bp1-EGFP knock-in ESCs showed, that these proteins not necessarily colocalize. This observation suggests the existence of a regulatory mechanism, coordinating M18bp1-Cenp-C interactions, potentially in a cell cycle-controlled way. It was demonstrated in *Moree et al. 2011*, that Cenp-C is necessary for centromeric Cenp-A targeting in *Xenopus* egg extracts. Accordingly, we also detected reduced Cenp-A levels after knock-down of Cenp-C in mouse ESCs.

M18bp1 is important for Hjurp-dependent Cenp-A deposition (Dunleavy et al. 2009; Foltz et al. 2009). We argue, that Cenp-C facilitates the recruitment of M18bp1 to centromeres, which is a critical step for subsequent Hjurp-mediated Cenp-A integration. Thus, depletion of Cenp-C eventually results in defective Cenp-A deposition. Nevertheless open questions remain concerning the recruitment of Cenp-C and M18bp1. Since M18bp1 could be identified as a H3K9me3 binder in peptide pull-down experiments (unpublished data, Silvia Dambacher), it will be of substantial interest, to analyze if H3K9me3 plays a role in M18bp1 recruitment to centromeric heterochromatin. Future studies on a potential interplay of pericentromeric and centromeric heterochromatin will further help to understand fundamental mechanistic principles in centromere formation.

3.5 Outlook

Taken together, this thesis provided novel findings on how centromeric and pericentromeric heterochromatin modifications and proteins participate in organismal development and cellular processes.

We investigated the role of Suv4-20h function in developmental differentiation. Given the fact, that Suv4-20h-dependent XOct-25 and Oct4 regulation is crucial for proper differentiation, it will be interesting to more closely investigate the phenotypes of (conditional) Suv4-20h knock-out mice. Some KO phenotypes, like e.g. class switch recombination defects in B cells (Schotta et al. 2008) or defects in lung epithelia, have already been observed in our lab. Gene expression profiling in these cells and tissues could reveal potential targets of Suv4-20h-mediated gene silencing that may contribute to the observed phenotypes.

The characterization of the regulatory network, modulating the novel heterochromatin modification H3K56me3, is an entrypoint for future analyses of its biological relevance. The human homolog of the mouse pseudogene of the KDM Jmjd2e is highly expressed in testis. It will be interesting to analyse, if and why the active removal of H3K9me3/H3K56me3 is important in this particular tissue. Furthermore, pull-down experiments with nucleosomes carrying the H3K56me3 modification may be useful to isolate interacting proteins. Knowing these modification “readers” will be valuable in identifying potential functions of H3K56me3.

In *Hahn et al.* we could demonstrate a severe impact of Suv4-20h depletion on organization of chromatin architecture and cohesin recruitment. To identify additional (tissue specific) interactors of Suv4-20h proteins, I generated knock-in mice

(unpublished data) expressing the endogenous Suv4-20h1 and Suv4-20h2 proteins with biochemical tags (HA-FLAG or V5-biotin) at their C-termini (Figure 7).

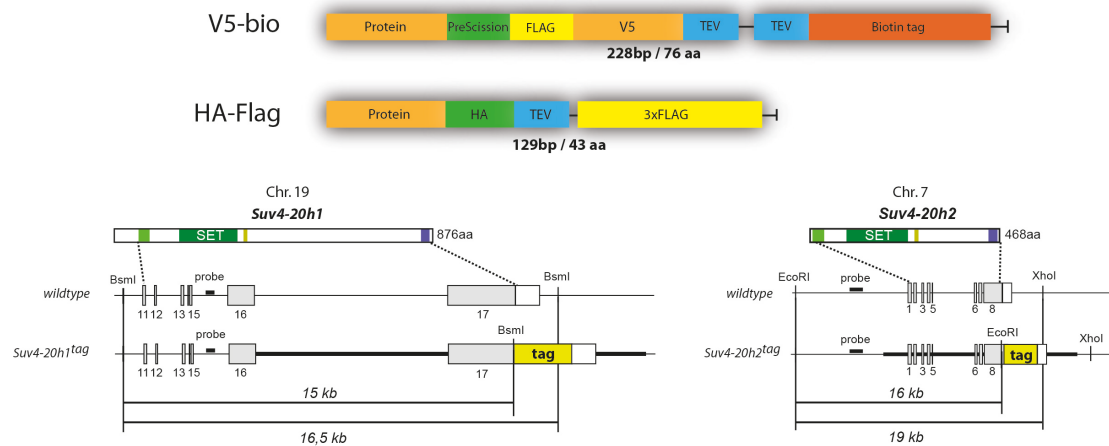


Figure 7. Suv4-20h knock-in strategy for protein tags

HA, Flag and V5 are standard affinity tags. The biotin-tag requires biotinylation by the BirA biotin ligase which is crossed into the corresponding mouse lines. PreScission protease and TEV are protease cleavage sites to elute proteins off beads. The knock-in strategies for the Suv4-20h loci are depicted in the lower part of the figure.

The HA-Flag tag is suited for tandem-affinity purification of protein complexes and was already successfully used in the *Hahn et al.* and in preliminary experiments, for immunoprecipitation of Suv4-20h1 proteins from knock-in mouse liver and brain tissues (data not shown). The V5-bio tag is a high-affinity protein tag and should be ideally suited for Chromatin Immunoprecipitation of the proteins (Kim et al. 2008). Suv4-20h1^{HA-FLAG}, Suv4-20h1^{V5-bio} and Suv4-20h2^{V5bio} knock-in mice are maintained in our lab and different mouse embryonic cell lines have been derived. These mice and cell lines, will be helpful in identifying tissue-specific genomic targets and interacting proteins, to characterize additional, yet unknown functions of Suv4-20h enzymes in cellular operations. Another noteworthy observation was, that nuclear pores are more evenly distributed in Suv4-20h knock-out cells. To gain more insight into this phenotype, we stained nuclear pores and performed high resolution microscopy in wild type and Suv4-20h mutant cells. As Lamin-B interacts with H4K20me2, we will further investigate Lamin A/B mobilities in Suv4-20h mutant cells. A study in a cellular model of progeria, a lamin-associated disease, showed severe upregulation of H4K20me3 (Shumaker et al. 2006) suggesting that Suv4-20h enzymes are important for nuclear envelope integrity. It will be interesting to elaborate on a potential mechanistic link between Suv4-20h enzymes, lamin arrangements and nuclear pore distributions.

Furthermore, we have established a powerful toolset for M18bp1 analyses, e.g. endogenous protein tagging, inducible knock-out cell lines and conditional knock-out mice. In preliminary experiments we could detect severe knock-out phenotypes in early embryogenesis and adult animals. Detailed analyses of these defects, will enhance our understanding of centromere establishment and help describing the complex M18bp1 interaction network.

Curriculum Vitae

Personal information

Name	Matthias Ulrich Hahn
Date of birth	17.12.1981
Place of birth	Ulm
Nationality	German

University education

02/2008 – 06/2013	PhD thesis and research projects in the group of Prof. Dr. Gunnar Schotta at the Adolf-Butenandt-Institute, LMU Munich
10/2007 – 12/2007	Research project in the group of Prof. Dr. Michael Boshart at the Institute of Genetics, LMU Munich
10/2006 – 09/2007	Diploma thesis in the group of Prof. Dr. Michael Boshart at the Institute of Genetics, LMU Munich: „Stadienspezifische Regulation der Enzyme des Citratzyklus in <i>T.brucei</i> “
08/2005 – 10/2005	Research project in the group of Dr. Frédéric Bringaud at the Bordeaux II Victor Segalen University, France
10/2002 – 09/2007	Study of biology at the LMU Munich, Germany (Diploma) Subjects: genetics, cell biology, biochemistry

School education

09/1996 - 06/2001	Abitur at Gymnasium Bad Aibling, Germany
-------------------	--

Acknowledgements

I thank Prof. Dr. Gunnar Schotta, for the opportunity to conduct my PhD research in his lab, for the intense discussions and his ideas that contributed to this thesis.

I thank Prof. Dr. Peter Becker for the opportunity to work in the inspiring atmosphere of the Adolf-Butenandt-Institute and for the supervision of my graduation at the biological faculty of the LMU.

I also want to thank my thesis advisory committee, Prof. Dr. Heinrich Leonhardt and PD Dr. Sandra Hake for fruitful project discussions and suggestions.

Furthermore I would like to thank all authors and co-authors of the publications mentioned on page 7, for contributing to the success of this thesis.

Next, I thank the current and past members of the Schotta group for having an interesting and fun time in the lab. Without the „*Great Sadic's*“ magic in-fusion cloning hands, a lot of knock-down experiments and transgenic cell-lines would have never seen the light of day. „*Professor*“ Nuber's image analysis qualities and Maike Hofmann's helping hands in the HP1 binding studies were substantial to finish our manuscript marathons in time.

Most of all, I thank my friends, family and parents - in particular my mother and my father (*in Memoriam* Ulrich Reinhard Hahn), for your constant support in everything I do.

References

- Agger K, Cloos PA, Christensen J, Pasini D, Rose S, Rappsilber J, Issaeva I, Canaani E, Salcini AE, Helin K. 2007. UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature* **449**: 731-734.
- Al-Haboubi T, Shumaker DK, Koser J, Wehnert M, Fahrenkrog B. 2011. Distinct association of the nuclear pore protein Nup153 with A- and B-type lamins. *Nucleus* **2**: 500-509.
- Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. 2007. High-resolution profiling of histone methylations in the human genome. *Cell* **129**: 823-837.
- Bergmann JH, Rodriguez MG, Martins NM, Kimura H, Kelly DA, Masumoto H, Larionov V, Jansen LE, Earnshaw WC. 2011. Epigenetic engineering shows H3K4me2 is required for HJURP targeting and CENP-A assembly on a synthetic human kinetochore. *The EMBO journal* **30**: 328-340.
- Bernard P, Maure JF, Partridge JF, Genier S, Javerzat JP, Allshire RC. 2001. Requirement of heterochromatin for cohesion at centromeres. *Science* **294**: 2539-2542.
- Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K et al. 2006. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**: 315-326.
- Bhattacharya D, Talwar S, Mazumder A, Shivashankar GV. 2009. Spatio-temporal plasticity in chromatin organization in mouse cell differentiation and during Drosophila embryogenesis. *Biophysical journal* **96**: 3832-3839.
- Black BE, Cleveland DW. 2011. Epigenetic centromere propagation and the nature of CENP-a nucleosomes. *Cell* **144**: 471-479.
- Boyer LA, Latek RR, Peterson CL. 2004. The SANT domain: a unique histone-tail-binding module? *Nature reviews Molecular cell biology* **5**: 158-163.
- Brown KE, Amoils S, Horn JM, Buckle VJ, Higgs DR, Merkenschlager M, Fisher AG. 2001. Expression of alpha- and beta-globin genes occurs within different nuclear domains in haemopoietic cells. *Nature cell biology* **3**: 602-606.
- Brown KE, Guest SS, Smale ST, Hahm K, Merkenschlager M, Fisher AG. 1997. Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. *Cell* **91**: 845-854.
- Bulut-Karslioglu A, Perrera V, Scaranaro M, de la Rosa-Velazquez IA, van de Nobelen S, Shukeir N, Popow J, Gerle B, Opravil S, Pagani M et al. 2012. A transcription factor-based mechanism for mouse heterochromatin formation. *Nature structural & molecular biology* **19**: 1023-1030.
- Burke B, Stewart CL. 2012. The nuclear lamins: flexibility in function. *Nature reviews Molecular cell biology*.
- Catena R, Ronfani L, Sassone-Corsi P, Davidson I. 2006. Changes in intranuclear chromatin architecture induce bipolar nuclear localization of histone variant H1T2 in male haploid spermatids. *Developmental biology* **296**: 231-238.
- Celeste A, Petersen S, Romanienko PJ, Fernandez-Capetillo O, Chen HT, Sedelnikova OA, Reina-San-Martin B, Coppola V, Meffre E, Difilippantonio MJ et al. 2002. Genomic instability in mice lacking histone H2AX. *Science* **296**: 922-927.
- Celic I, Masumoto H, Griffith WP, Meluh P, Cotter RJ, Boeke JD, Verreault A. 2006. The sirtuins hst3 and Hst4p preserve genome integrity by controlling histone h3 lysine 56 deacetylation. *Current biology : CB* **16**: 1280-1289.
- Cheutin T, McNairn AJ, Jenuwein T, Gilbert DM, Singh PB, Misteli T. 2003. Maintenance of stable heterochromatin domains by dynamic HP1 binding. *Science* **299**: 721-725.

- Dambacher S, Hahn M, Schotta G. 2010. Epigenetic regulation of development by histone lysine methylation. *Heredity* **105**: 24-37.
- Daujat S, Weiss T, Mohn F, Lange UC, Ziegler-Birling C, Zeissler U, Lappe M, Schubeler D, Torres-Padilla ME, Schneider R. 2009. H3K64 trimethylation marks heterochromatin and is dynamically remodeled during developmental reprogramming. *Nature structural & molecular biology* **16**: 777-781.
- Dialynas GK, Terjung S, Brown JP, Aucott RL, Baron-Luhr B, Singh PB, Georgatos SD. 2007. Plasticity of HP1 proteins in mammalian cells. *Journal of cell science* **120**: 3415-3424.
- Driscoll R, Hudson A, Jackson SP. 2007. Yeast Rtt109 promotes genome stability by acetylating histone H3 on lysine 56. *Science* **315**: 649-652.
- Dunleavy EM, Roche D, Tagami H, Lacoste N, Ray-Gallet D, Nakamura Y, Daigo Y, Nakatani Y, Almouzni-Pettinotti G. 2009. HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres. *Cell* **137**: 485-497.
- Eckert CA, Gravidahl DJ, Megee PC. 2007. The enhancement of pericentromeric cohesin association by conserved kinetochore components promotes high-fidelity chromosome segregation and is sensitive to microtubule-based tension. *Genes & development* **21**: 278-291.
- Erdel F, Schubert T, Marth C, Langst G, Rippe K. 2010. Human ISWI chromatin-remodeling complexes sample nucleosomes via transient binding reactions and become immobilized at active sites. *Proceedings of the National Academy of Sciences of the United States of America* **107**: 19873-19878.
- Eymery A, Callanan M, Vourc'h C. 2009. The secret message of heterochromatin: new insights into the mechanisms and function of centromeric and pericentric repeat sequence transcription. *The International journal of developmental biology* **53**: 259-268.
- Fan Y, Nikitina T, Zhao J, Fleury TJ, Bhattacharyya R, Bouhassira EE, Stein A, Woodcock CL, Skoultschi AI. 2005. Histone H1 depletion in mammals alters global chromatin structure but causes specific changes in gene regulation. *Cell* **123**: 1199-1212.
- Feldman N, Gerson A, Fang J, Li E, Zhang Y, Shinkai Y, Cedar H, Bergman Y. 2006. G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. *Nature cell biology* **8**: 188-194.
- Ferraiuolo MA, Rousseau M, Miyamoto C, Shenker S, Wang XQ, Nadler M, Blanchette M, Dostie J. 2010. The three-dimensional architecture of Hox cluster silencing. *Nucleic acids research* **38**: 7472-7484.
- Fodor BD, Kubicek S, Yonezawa M, O'Sullivan RJ, Sengupta R, Perez-Burgos L, Opravil S, Mechtler K, Schotta G, Jenuwein T. 2006. Jmjd2b antagonizes H3K9 trimethylation at pericentric heterochromatin in mammalian cells. *Genes & development* **20**: 1557-1562.
- Foltz DR, Jansen LE, Bailey AO, Yates JR, 3rd, Bassett EA, Wood S, Black BE, Cleveland DW. 2009. Centromere-specific assembly of CENP-a nucleosomes is mediated by HJURP. *Cell* **137**: 472-484.
- Fraga MF, Ballestar E, Villar-Garea A, Boix-Chornet M, Espada J, Schotta G, Bonaldi T, Haydon C, Ropero S, Petrie K et al. 2005. Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nature genetics* **37**: 391-400.
- Fujita Y, Hayashi T, Kiyomitsu T, Toyoda Y, Kokubu A, Obuse C, Yanagida M. 2007. Priming of centromere for CENP-A recruitment by human hMis18alpha, hMis18beta, and M18BP1. *Developmental cell* **12**: 17-30.
- Garcia BA, Hake SB, Diaz RL, Kauer M, Morris SA, Recht J, Shabanowitz J, Mishra N, Strahl BD, Allis CD et al. 2007. Organismal differences in post-translational

- modifications in histones H3 and H4. *The Journal of biological chemistry* **282**: 7641-7655.
- Gondor A, Ohlsson R. 2009. Chromosome crosstalk in three dimensions. *Nature* **461**: 212-217.
- Gonzalo S, Garcia-Cao M, Fraga MF, Schotta G, Peters AH, Cotter SE, Eguia R, Dean DC, Esteller M, Jenuwein T et al. 2005. Role of the RB1 family in stabilizing histone methylation at constitutive heterochromatin. *Nature cell biology* **7**: 420-428.
- Grewal SI, Jia S. 2007. Heterochromatin revisited. *Nature reviews Genetics* **8**: 35-46.
- Guenatri M, Bailly D, Maison C, Almouzni G. 2004. Mouse centric and pericentric satellite repeats form distinct functional heterochromatin. *The Journal of cell biology* **166**: 493-505.
- Hahn M, Dambacher S, Schotta G. 2010. Heterochromatin dysregulation in human diseases. *Journal of applied physiology* **109**: 232-242.
- Hancock R. 2004. A role for macromolecular crowding effects in the assembly and function of compartments in the nucleus. *Journal of structural biology* **146**: 281-290.
- Harrington JJ, Van Bokkelen G, Mays RW, Gustashaw K, Willard HF. 1997. Formation of de novo centromeres and construction of first-generation human artificial microchromosomes. *Nature genetics* **15**: 345-355.
- Hayashi T, Fujita Y, Iwasaki O, Adachi Y, Takahashi K, Yanagida M. 2004. Mis16 and Mis18 are required for CENP-A loading and histone deacetylation at centromeres. *Cell* **118**: 715-729.
- Henikoff S, Ahmad K, Malik HS. 2001. The centromere paradox: stable inheritance with rapidly evolving DNA. *Science* **293**: 1098-1102.
- Hirano Y, Hizume K, Kimura H, Takeyasu K, Haraguchi T, Hiraoka Y. 2012. Lamin B receptor recognizes specific modifications of histone H4 in heterochromatin formation. *The Journal of biological chemistry*.
- Jansen LE, Black BE, Foltz DR, Cleveland DW. 2007. Propagation of centromeric chromatin requires exit from mitosis. *The Journal of cell biology* **176**: 795-805.
- Joffe B, Leonhardt H, Solovei I. 2010. Differentiation and large scale spatial organization of the genome. *Current opinion in genetics & development* **20**: 562-569.
- Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, Ebmeier CC, Goossens J, Rahl PB, Levine SS et al. 2010. Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**: 430-435.
- Kapoor M, Montes de Oca Luna R, Liu G, Lozano G, Cummings C, Mancini M, Ouspenski I, Brinkley BR, May GS. 1998. The cenpB gene is not essential in mice. *Chromosoma* **107**: 570-576.
- Kapoor-Vazirani P, Kagey JD, Vertino PM. 2011. SUV420H2-mediated H4K20 trimethylation enforces RNA polymerase II promoter-proximal pausing by blocking hMOF-dependent H4K16 acetylation. *Molecular and cellular biology* **31**: 1594-1609.
- Keller GM. 1995. In vitro differentiation of embryonic stem cells. *Current opinion in cell biology* **7**: 862-869.
- Koch B, Kueng S, Ruckenbauer C, Wendt KS, Peters JM. 2008. The Suv39h-HP1 histone methylation pathway is dispensable for enrichment and protection of cohesin at centromeres in mammalian cells. *Chromosoma* **117**: 199-210.
- Korber P, Luckenbach T, Blaschke D, Horz W. 2004. Evidence for histone eviction in trans upon induction of the yeast PHO5 promoter. *Molecular and cellular biology* **24**: 10965-10974.
- Lachner M, Sengupta R, Schotta G, Jenuwein T. 2004. Trilogies of histone lysine methylation as epigenetic landmarks of the eukaryotic genome. *Cold Spring Harbor symposia on quantitative biology* **69**: 209-218.

- Lan F, Bayliss PE, Rinn JL, Whetstone JR, Wang JK, Chen S, Iwase S, Alpatov R, Issaeva I, Canaani E et al. 2007. A histone H3 lysine 27 demethylase regulates animal posterior development. *Nature* **449**: 689-694.
- Lee TI, Jenner RG, Boyer LA, Guenther MG, Levine SS, Kumar RM, Chevalier B, Johnstone SE, Cole MF, Isono K et al. 2006. Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* **125**: 301-313.
- Lehnertz B, Ueda Y, Derijck AA, Braunschweig U, Perez-Burgos L, Kubicek S, Chen T, Li E, Jenuwein T, Peters AH. 2003. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Current biology : CB* **13**: 1192-1200.
- Li JY, Pu MT, Hirasawa R, Li BZ, Huang YN, Zeng R, Jing NH, Chen T, Li E, Sasaki H et al. 2007. Synergistic function of DNA methyltransferases Dnmt3a and Dnmt3b in the methylation of Oct4 and Nanog. *Molecular and cellular biology* **27**: 8748-8759.
- Loyola A, Tagami H, Bonaldi T, Roche D, Quivy JP, Imhof A, Nakatani Y, Dent SY, Almouzni G. 2009. The HP1alpha-CAF1-SetDB1-containing complex provides H3K9me1 for Suv39-mediated K9me3 in pericentric heterochromatin. *EMBO reports* **10**: 769-775.
- Lu X, Hamkalo B, Parseghian MH, Hansen JC. 2009a. Chromatin condensing functions of the linker histone C-terminal domain are mediated by specific amino acid composition and intrinsic protein disorder. *Biochemistry* **48**: 164-172.
- Lu X, Wontakal SN, Emelyanov AV, Morcillo P, Konev AY, Fyodorov DV, Skoultschi AI. 2009b. Linker histone H1 is essential for Drosophila development, the establishment of pericentric heterochromatin, and a normal polytene chromosome structure. *Genes & development* **23**: 452-465.
- Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**: 251-260.
- Maddox PS, Hyndman F, Monen J, Oegema K, Desai A. 2007. Functional genomics identifies a Myb domain-containing protein family required for assembly of CENP-A chromatin. *The Journal of cell biology* **176**: 757-763.
- Magklara A, Yen A, Colquitt BM, Clowney EJ, Allen W, Markenscoff-Papadimitriou E, Evans ZA, Kheradpour P, Mountoufaris G, Carey C et al. 2011. An epigenetic signature for monoallelic olfactory receptor expression. *Cell* **145**: 555-570.
- Martens JH, O'Sullivan RJ, Braunschweig U, Opravil S, Radolf M, Steinlein P, Jenuwein T. 2005. The profile of repeat-associated histone lysine methylation states in the mouse epigenome. *The EMBO journal* **24**: 800-812.
- Masumoto H, Hawke D, Kobayashi R, Verreault A. 2005. A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response. *Nature* **436**: 294-298.
- Mayer R, Brero A, von Hase J, Schroeder T, Cremer T, Dietzel S. 2005. Common themes and cell type specific variations of higher order chromatin arrangements in the mouse. *BMC cell biology* **6**: 44.
- Meshorer E, Misteli T. 2006. Chromatin in pluripotent embryonic stem cells and differentiation. *Nature reviews Molecular cell biology* **7**: 540-546.
- Meshorer E, Yellajoshula D, George E, Scambler PJ, Brown DT, Misteli T. 2006. Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Developmental cell* **10**: 105-116.
- Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP et al. 2007. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**: 553-560.

- Mito Y, Henikoff JG, Henikoff S. 2005. Genome-scale profiling of histone H3.3 replacement patterns. *Nature genetics* **37**: 1090-1097.
- Moree B, Meyer CB, Fuller CJ, Straight AF. 2011. CENP-C recruits M18BP1 to centromeres to promote CENP-A chromatin assembly. *The Journal of cell biology* **194**: 855-871.
- Niwa H. 2007. How is pluripotency determined and maintained? *Development* **134**: 635-646.
- Nonaka N, Kitajima T, Yokobayashi S, Xiao G, Yamamoto M, Grewal SI, Watanabe Y. 2002. Recruitment of cohesin to heterochromatic regions by Swi6/HP1 in fission yeast. *Nature cell biology* **4**: 89-93.
- Olszak AM, van Essen D, Pereira AJ, Diehl S, Manke T, Maiato H, Sacconi S, Heun P. 2011. Heterochromatin boundaries are hotspots for de novo kinetochore formation. *Nature cell biology* **13**: 799-808.
- Onn I, Heidinger-Pauli JM, Guacci V, Unal E, Koshland DE. 2008. Sister chromatid cohesion: a simple concept with a complex reality. *Annual review of cell and developmental biology* **24**: 105-129.
- Pan G, Tian S, Nie J, Yang C, Ruotti V, Wei H, Jonsdottir GA, Stewart R, Thomson JA. 2007. Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. *Cell stem cell* **1**: 299-312.
- Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, Gregson HC, Jarmuz A, Canzonetta C, Webster Z, Nesterova T et al. 2008. Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* **132**: 422-433.
- Pedersen MT, Helin K. 2010. Histone demethylases in development and disease. *Trends in cell biology* **20**: 662-671.
- Perpelescu M, Fukagawa T. 2011. The ABCs of CENPs. *Chromosoma* **120**: 425-446.
- Peters AH, O'Carroll D, Scherthan H, Mechtler K, Sauer S, Schofer C, Weipoltshammer K, Pagani M, Lachner M, Kohlmaier A et al. 2001. Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* **107**: 323-337.
- Pinheiro I, Margueron R, Shukeir N, Eisold M, Fritzsche C, Richter FM, Mittler G, Genoud C, Goyama S, Kurokawa M et al. 2012. Prdm3 and Prdm16 are H3K9me1 methyltransferases required for mammalian heterochromatin integrity. *Cell* **150**: 948-960.
- Probst AV, Okamoto I, Casanova M, El Marjou F, Le Baccon P, Almouzni G. 2010. A strand-specific burst in transcription of pericentric satellites is required for chromocenter formation and early mouse development. *Developmental cell* **19**: 625-638.
- Rea S, Eisenhaber F, O'Carroll D, Strahl BD, Sun ZW, Schmid M, Opravil S, Mechtler K, Ponting CP, Allis CD et al. 2000. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406**: 593-599.
- Recht J, Tsubota T, Tanny JC, Diaz RL, Berger JM, Zhang X, Garcia BA, Shabanowitz J, Burlingame AL, Hunt DF et al. 2006. Histone chaperone Asf1 is required for histone H3 lysine 56 acetylation, a modification associated with S phase in mitosis and meiosis. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 6988-6993.
- Regnier V, Vagnarelli P, Fukagawa T, Zerjal T, Burns E, Trouche D, Earnshaw W, Brown W. 2005. CENP-A is required for accurate chromosome segregation and sustained kinetochore association of BubR1. *Molecular and cellular biology* **25**: 3967-3981.
- Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, Goodnough LH, Helms JA, Farnham PJ, Segal E et al. 2007. Functional demarcation of active

- and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **129**: 1311-1323.
- Rufiange A, Jacques PE, Bhat W, Robert F, Nourani A. 2007. Genome-wide replication-independent histone H3 exchange occurs predominantly at promoters and implicates H3 K56 acetylation and Asf1. *Molecular cell* **27**: 393-405.
- Sasai N, Defossez PA. 2009. Many paths to one goal? The proteins that recognize methylated DNA in eukaryotes. *The International journal of developmental biology* **53**: 323-334.
- Schneider AC, Heukamp LC, Rogenhofer S, Fechner G, Bastian PJ, von Ruecker A, Muller SC, Ellinger J. 2011. Global histone H4K20 trimethylation predicts cancer-specific survival in patients with muscle-invasive bladder cancer. *BJU international* **108**: E290-296.
- Schotta G, Lachner M, Sarma K, Ebert A, Sengupta R, Reuter G, Reinberg D, Jenuwein T. 2004. A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes & development* **18**: 1251-1262.
- Schotta G, Sengupta R, Kubicek S, Malin S, Kauer M, Callen E, Celeste A, Pagani M, Opravil S, De La Rosa-Velazquez IA et al. 2008. A chromatin-wide transition to H4K20 monomethylation impairs genome integrity and programmed DNA rearrangements in the mouse. *Genes & development* **22**: 2048-2061.
- Schuettengruber B, Martinez AM, Iovino N, Cavalli G. 2011. Trithorax group proteins: switching genes on and keeping them active. *Nature reviews Molecular cell biology* **12**: 799-814.
- Screpanti E, De Antoni A, Alushin GM, Petrovic A, Melis T, Nogales E, Musacchio A. 2011. Direct binding of Cenp-C to the Mis12 complex joins the inner and outer kinetochore. *Current biology : CB* **21**: 391-398.
- Shimura M, Toyoda Y, Iijima K, Kinomoto M, Tokunaga K, Yoda K, Yanagida M, Sata T, Ishizaka Y. 2011. Epigenetic displacement of HP1 from heterochromatin by HIV-1 Vpr causes premature sister chromatid separation. *The Journal of cell biology* **194**: 721-735.
- Shumaker DK, Dechat T, Kohlmaier A, Adam SA, Bozovsky MR, Erdos MR, Eriksson M, Goldman AE, Khun S, Collins FS et al. 2006. Mutant nuclear lamin A leads to progressive alterations of epigenetic control in premature aging. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 8703-8708.
- Silva MC, Bodor DL, Stellfox ME, Martins NM, Hochegger H, Foltz DR, Jansen LE. 2012. Cdk activity couples epigenetic centromere inheritance to cell cycle progression. *Developmental cell* **22**: 52-63.
- Singleton MK, Gonzales ML, Leung KN, Yasui DH, Schroeder DI, Dunaway K, LaSalle JM. 2011. MeCP2 is required for global heterochromatic and nucleolar changes during activity-dependent neuronal maturation. *Neurobiology of disease* **43**: 190-200.
- Souza PP, Volkel P, Trinel D, Vandamme J, Rosnoblet C, Heliot L, Angrand PO. 2009. The histone methyltransferase SUV420H2 and Heterochromatin Proteins HP1 interact but show different dynamic behaviours. *BMC cell biology* **10**: 41.
- Tachibana M, Sugimoto K, Nozaki M, Ueda J, Ohta T, Ohki M, Fukuda M, Takeda N, Niida H, Kato H et al. 2002. G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes & development* **16**: 1779-1791.
- Tan M, Luo H, Lee S, Jin F, Yang JS, Montellier E, Buchou T, Cheng Z, Rousseaux S, Rajagopal N et al. 2011. Identification of 67 histone marks and histone

- lysine crotonylation as a new type of histone modification. *Cell* **146**: 1016-1028.
- Terranova R, Sauer S, Merckenschlager M, Fisher AG. 2005. The reorganisation of constitutive heterochromatin in differentiating muscle requires HDAC activity. *Experimental cell research* **310**: 344-356.
- Thompson SL, Bakhoum SF, Compton DA. 2010. Mechanisms of chromosomal instability. *Current biology : CB* **20**: R285-295.
- Thomson M, Liu SJ, Zou LN, Smith Z, Meissner A, Ramanathan S. 2011. Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. *Cell* **145**: 875-889.
- Trojer P, Reinberg D. 2007. Facultative heterochromatin: is there a distinctive molecular signature? *Molecular cell* **28**: 1-13.
- Tryndyak VP, Kovalchuk O, Pogribny IP. 2006. Loss of DNA methylation and histone H4 lysine 20 trimethylation in human breast cancer cells is associated with aberrant expression of DNA methyltransferase 1, Suv4-20h2 histone methyltransferase and methyl-binding proteins. *Cancer biology & therapy* **5**: 65-70.
- Van Den Broeck A, Brambilla E, Moro-Sibilot D, Lantuejoul S, Brambilla C, Eymin B, Khochbin S, Gazzeri S. 2008. Loss of histone H4K20 trimethylation occurs in preneoplasia and influences prognosis of non-small cell lung cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research* **14**: 7237-7245.
- Varga-Weisz PD, Becker PB. 2006. Regulation of higher-order chromatin structures by nucleosome-remodelling factors. *Current opinion in genetics & development* **16**: 151-156.
- Vempati RK, Jayani RS, Notani D, Sengupta A, Galande S, Halder D. 2010. p300-mediated acetylation of histone H3 lysine 56 functions in DNA damage response in mammals. *The Journal of biological chemistry* **285**: 28553-28564.
- Vissel B, Choo KH. 1989. Mouse major (gamma) satellite DNA is highly conserved and organized into extremely long tandem arrays: implications for recombination between nonhomologous chromosomes. *Genomics* **5**: 407-414.
- Voullaire LE, Slater HR, Petrovic V, Choo KH. 1993. A functional marker centromere with no detectable alpha-satellite, satellite III, or CENP-B protein: activation of a latent centromere? *American journal of human genetics* **52**: 1153-1163.
- Williams SK, Truong D, Tyler JK. 2008. Acetylation in the globular core of histone H3 on lysine-56 promotes chromatin disassembly during transcriptional activation. *Proceedings of the National Academy of Sciences of the United States of America* **105**: 9000-9005.
- Wongtawan T, Taylor JE, Lawson KA, Wilmut I, Pennings S. 2011. Histone H4K20me3 and HP1alpha are late heterochromatin markers in development, but present in undifferentiated embryonic stem cells. *Journal of cell science* **124**: 1878-1890.
- Wurtele H, Kaiser GS, Bacal J, St-Hilaire E, Lee EH, Tsao S, Dorn J, Maddox P, Lisby M, Pasero P et al. 2012. Histone H3 lysine 56 acetylation and the response to DNA replication fork damage. *Molecular and cellular biology* **32**: 154-172.
- Yu Y, Song C, Zhang Q, DiMaggio PA, Garcia BA, York A, Carey MF, Grunstein M. 2012. Histone H3 lysine 56 methylation regulates DNA replication through its interaction with PCNA. *Molecular cell* **46**: 7-17.
- Zhao J, Sun BK, Erwin JA, Song JJ, Lee JT. 2008. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* **322**: 750-756.